Stochastik I

Notizen* zu einer Vorlesung an der
Johannes-Gutenberg-Universität Mainz, Sommer 2017

Matthias Birkner

Vorläufige Version, 20. Juni 2018

Kommentare, Korrekturvorschläge, Hinweise auf (Tipp-)fehler gerne per Email an `birkner@mathematik.uni-mainz.de` senden

*Diese Notizen ersetzen – zumindest nach der Meinung des Autors – nicht den Besuch der Vorlesung und auch nicht die Nachbereitung anhand eigener Notizen und der Literatur. Zudem ist für das Verständnis des Stoffes der Besuch der Übungsgruppe und eigenständiges Bearbeiten der Übungsaufgaben überaus zu empfehlen.

Inhaltsverzeichnis

Ich danke Melanie Behnke, die durch ihre kritische Lektüre dieser Notizen an verschiedenen Stellen zur Korrektur von Ungenauigkeiten und Tippfehlern beigetragen hat.

Kapitel 1

Grundlagen der Maßtheorie

In diesem Kapitel geht es um grundlegende Begriffe und Sätze der Maßtheorie, insoweit sie für die Diskussion und Fundierung der Stochastik benötigt werden. Nachzulesen (und z.T. weiterführend) z.B. bei Klenke [Kl, Kapitel 1], Williams [Wi, Chapter 1], Kallenberg [Ka, Chapter 1 and 2], Durrett [Du, Appendix].

1.1 Mengensysteme, σ-Algebren

Sei Ω (nicht-leere) Menge (der „Ergebnisraum“ oder „Stichprobenraum“) $2^\Omega := \{B : B \subset \Omega\}$ die Potenzmenge.

Definition 1.1. $\mathscr{A} \subset 2^\Omega$ heißt eine *σ-Algebra* (über Ω), falls gilt

$$\begin{aligned} &i) && \Omega \in \mathscr{A}, \\ &ii) && A \in \mathscr{A} \;\Rightarrow\; A^c \in \mathscr{A}, \\ &iii) && A_1, A_2, \dots \in \mathscr{A} \;\Rightarrow\; \bigcup_{n=1}^\infty A_n \in \mathscr{A}. \end{aligned}$$

Falls $\mathscr{A}$ *i)*, *ii)* und

$$iii') \quad A, B \in \mathscr{A} \;\Rightarrow\; A \cap B \in \mathscr{A}$$

erfüllt, so heißt $\mathscr{A}$ eine *Algebra*.

Beobachtung 1.2. Ist $\mathscr{A}$ eine σ-Algebra, so gilt auch

$$\begin{aligned} &iv) && \varnothing = \Omega^c \in \mathscr{A}, \\ &v) && A_1, A_2, \dots \in \mathscr{A} \;\Rightarrow\; \bigcap_{n=1}^\infty A_n = \Big(\bigcup_{n=1}^\infty A_n^c\Big)^c \in \mathscr{A} \end{aligned}$$

(wir verwenden die de Morgan'sche Regel: $\big(\bigcap_n A_n\big)^c = \bigcup_n A_n^c$).

Eine Algebra $\mathscr{A}$ erfüllt

$$A_1, A_2, \dots, A_n \in \mathscr{A} \;\Rightarrow\; A_1 \cup A_2 \cup \dots \cup A_n \in \mathscr{A} \;\text{ und }\; A_1 \cap A_2 \cup \dots \cap A_n \in \mathscr{A}$$

$\{\varnothing, \Omega\}$ und 2^Ω sind σ-Algebren. Eine σ-Algebra ist insbesondere eine Algebra, denn

$$A \cup B = A \cup B \cup \varnothing \cup \varnothing \cup \varnothing \cup \cdots$$

Definition 1.3. Ein Paar $(\Omega, \mathscr{A})$ mit $\mathscr{A}$ σ-Algebra über Ω heißt ein *messbarer Raum* (auch: Messraum oder Ereignisraum).

$A \in \mathscr{A}$ heißt ($\mathscr{A}$-)messbare Teilmenge von Ω.

Bemerkung 1.4. Seien $\mathscr{A}_i$, $i \in I$ σ-Algebren über Ω (wobei I eine beliebige Indexmenge), so ist

$$\bigcap_{i \in I} \mathscr{A}_i \quad \text{ebenfalls eine } \sigma\text{-Algebra über } \Omega.$$

Beweis. i), ii) sind klar, zu *iii)*:

$$A_1, A_2, \dots \in \bigcap_{i \in I} \mathscr{A}_i \iff \forall\, n \in \mathbb{N}, i \in I \,:\, A_n \in \mathscr{A}_i,$$

somit gilt

$$A_1, A_2, \dots \in \bigcap_{i \in I} \mathscr{A}_i \quad \Rightarrow \quad \forall\, i \in I \,:\, \bigcup_n A_n \in \mathscr{A}_i \quad \Rightarrow \quad \bigcup_n A_n \in \bigcap_{i \in I} \mathscr{A}_i.$$

□

Definition 1.5. Für $\mathscr{C} \subset 2^\Omega$ heißt

$$\sigma(\mathscr{C}) := \bigcap \left\{ \mathscr{A} : \mathscr{A} \text{ ist } \sigma\text{-Algebra mit } \mathscr{C} \subset \mathscr{A} \subset 2^\Omega \right\},$$

die von $\mathscr{C}$ erzeugte σ-Algebra. Dies ist offenbar die kleinste σ-Algebra über Ω, die $\mathscr{C}$ umfasst.

Erinnerung 1.6 (Topologie)**.** $\tau \subset 2^\Omega$ heißt eine Topologie (auf Ω), wenn gilt

i) $\varnothing, \Omega \in \tau$,

ii) $A, B \in \tau \Rightarrow A \cap B \in \tau$

iii) $\mathcal{F} \subset \tau$ beliebige Teilmenge $\Rightarrow \bigcup_{A \in \mathcal{F}} A \in \tau$.

$A \in \tau$ sind/heißen die offenen Teilmengen in Ω (bzgl. τ).

Wenn Ω eine Metrik d trägt (d.h. $d : \Omega \times \Omega \to [0, \infty)$ mit $\forall\, x, y, z \in \Omega$: i) $d(x, y) = d(y, x)$, ii) $d(x, y) = 0 \iff x = y$, iii) $d(x, z) \le d(x, y) + d(y, z)$),

so erzeugt d eine Topologie τ via

$$A \in \tau \iff \forall\, x \in A : \exists\, r > 0 : B_r(x) \subset A$$

wo $B_r(x) := \{y \in \Omega : d(x, y) < r\}$ die offene Kugel um x mit Radius r ist.

Auf $\Omega = \mathbb{R}^m$ verwenden wir (meist)

$$d(x, y) = \Big(\sum_{i=1}^m (x_i - y_i)^2 \Big)^{1/2}, \quad x = (x_1, \dots, x_m), y = (y_1, \dots, y_m),$$

den Euklidischen Abstand.

Definition 1.7. Wenn Ω eine Topologie τ trägt, so heißt die von den offenen Mengen erzeugte σ-Algebra $\sigma(\tau)$ die *Borel-σ-Algebra*[1], man schreibt $\mathcal{B}(\Omega) := \sigma(\tau)$.

Beispiel 1.8 (Borel-σ-Algebra auf $\mathbb{R}^d$)**.** Für $a = (a_1, \dots, a_d), b = (b_1, \dots, b_d) \in \mathbb{R}^d$ schreibe $a \leq b$ (bzw. $a < b$), wenn $a_i \leq b_i$ (bzw. $a_i < b_i$) für $i = 1, 2, \dots, d$,

$$[a,b] := \{x \in \mathbb{R}^d : a \leq x \leq b\} = \bigtimes_{i=1}^{d} [a_i, b_i]$$

analog $(a,b]$, (a,b).

Jedes der folgenden Mengensysteme erzeugt $\mathcal{B}(\mathbb{R}^d)$:

$$\begin{aligned}
\mathscr{C}_0 &:= \left\{B_r(x) : x \in \mathbb{Q}^d, r \in \mathbb{Q} \cap (0,\infty)\right\} \\
\mathscr{C}_1 &:= \left\{(-\infty, b] : b \in \mathbb{R}^d\right\} \\
\mathscr{C}_2 &:= \left\{(a, b] : a, b \in \mathbb{R}^d, a < b\right\} \\
\mathscr{C}_3 &:= \left\{[a, b] : a, b \in \mathbb{R}^d, a \leq b\right\} \\
\mathscr{C}_4 &:= \left\{(a, b) : a, b \in \mathbb{R}^d, a < b\right\}
\end{aligned}$$

Beweis. Es genügt für $i = 0, \dots, 4$ zu zeigen, dass $\mathscr{C}_i \subset \mathcal{B}(\mathbb{R}^d)$ und dass $\mathcal{O} \in \sigma(\mathscr{C}_i)$ für jede offene Menge $\mathcal{O} \subset \mathbb{R}^d$ (denn $\mathscr{C}_i \subset \mathcal{B}(\mathbb{R}^d) \Rightarrow \sigma(\mathscr{C}_i) \subset \mathcal{B}(\mathbb{R}^d)$; $\tau \subset \sigma(\mathscr{C}_i) \Rightarrow \sigma(\tau) \subset \sigma(\mathscr{C}_i)$).

Betrachte $\mathscr{C}_0$: Offenbar ist $\mathscr{C}_0 \subset \tau \subset \mathcal{B}(\mathbb{R}^d)$, $\mathcal{O} \subset \mathbb{R}^d$ offen lässt sich darstellen als

$$\mathcal{O} = \bigcup_{n=1}^{\infty} B_{r_n}(x_n) \in \sigma(\mathscr{C}_0)$$

mit geeigneten $x_n \in \mathbb{Q}^d \cap \mathcal{O}$ und $r_n \in \mathbb{Q}_+$. $\mathscr{C}_4$ kann analog behandelt werden. Wegen

$$\begin{aligned}
[a,b] &= \bigcap_n \left(a - (\tfrac{1}{n}, \dots, \tfrac{1}{n}), b + (\tfrac{1}{n}, \dots, \tfrac{1}{n})\right) \in \sigma(\mathscr{C}_4), \\
(a,b) &= \bigcup_n \left[a + (\tfrac{1}{n}, \dots, \tfrac{1}{n}), b - (\tfrac{1}{n}, \dots, \tfrac{1}{n})\right] \in \sigma(\mathscr{C}_3)
\end{aligned}$$

ist $\sigma(\mathscr{C}_3) = \sigma(\mathscr{C}_4)$, analog $\sigma(\mathscr{C}_2) = \sigma(\mathscr{C}_4)$.

Wegen

$$(-\infty, b] = \bigcup_n \left[(-n, \dots, -n), b\right] \in \sigma(\mathscr{C}_3)$$

ist $\sigma(\mathscr{C}_1) \subset \sigma(\mathscr{C}_3)$, wegen

$$(a,b] = (-\infty, b] \setminus \left(\bigcup_{i=1}^{d} (-\infty, c^{(i)}]\right) \quad \text{mit } c^{(i)} = (b_1, \dots, b_{i-1}, a_i, b_{i+1}, \dots, b_d)$$

ist $\sigma(\mathscr{C}_2) \subset \sigma(\mathscr{C}_1)$. $\square$

[1] nach Émile Borel, 1871–1956

Bemerkung. Für $\Omega = \mathbb{R}$ ist

$$\mathscr{A} = \Big\{ \bigcup_{j=1}^{k} I_j \ : \ k \in \mathbb{N}, I_1, \dots, I_k \text{ paarw. disjunkte (möglicherw. halb-unendliche) Intervalle} \Big\}$$

eine Algebra.

Definition 1.9. $\mathscr{C} \subset 2^\Omega$ heißt *schnittstabil* , wenn

$$\text{für alle } \ A, B \in \mathscr{C} : \ A \cap B \in \mathscr{C}.$$

Beobachtung. Die Mengensysteme $\mathscr{C}_1 - \mathscr{C}_4$ aus Bsp. 1.8 sind $\cap$-stabil.

Definition 1.10. $\Omega \neq \varnothing$. $\mathscr{D} \subset 2^\Omega$ heißt ein *Dynkin-System* (auch: ein λ-System), wenn gilt

i) $\Omega \in \mathscr{D}$,

ii) $A \in \mathscr{D} \ \Rightarrow \ A^c \in \mathscr{D}$,

iii) $A_1, A_2, \dots \in \mathscr{D}$ paarweise disjunkt $\ \Rightarrow \ \dot{\bigcup}_{n=1}^{\infty} A_n \in \mathscr{D}$.

Bemerkung 1.11.

1. $\mathscr{A}$ σ-Algebra $\Rightarrow$ $\mathscr{A}$ Dynkin-System
2. $\mathscr{D}$ Dynkin-System, so ist $\varnothing = \Omega^c \in \mathscr{D}$, für $D_1 \dots, D_n \in \mathscr{D}$ paarweise disjunkt ist $D_1 \dot\cup D_2 \dot\cup \cdots \dot\cup D_n = D_1 \dot\cup D_2 \dot\cup \cdots \cup D_n \dot\cup \varnothing \dot\cup \varnothing \dot\cup \cdots \in \mathscr{D}$.
3. $\mathscr{D}$ Dynkin-System, $D_1, D_2 \in \mathscr{D}$ mit $D_1 \subset D_2$, so ist $D_2 \setminus D_1 = (D_1 \dot\cup D_2^c)^c \in \mathscr{D}$.
4. Im Allgemeinen ist ein Dynkin-System nicht abgeschlossen unter $\cap$ (und somit keine Algebra). Betrachte z.B. Ω endlich mit $|\Omega| \in 2\mathbb{N}$, $\mathscr{D} := \{A \subset \Omega : |A| \text{ ist gerade}\}$ ist ein Dynkin-System.

Proposition 1.12. *Ein Dynkin-System $\mathscr{D}$ ist eine σ-Algebra g.d.w. $\mathscr{D}$ $\cap$-stabil ist.*

Beweis. „$\Rightarrow$“ : ✓

„$\Leftarrow$“ : Seien $A_1, A_2, \dots \in \mathscr{D}$. Wir konstruieren (induktiv) $(D_n)_{n\in\mathbb{N}} \subset \mathscr{D}$, D_i paarweise disjunkt mit $\dot\bigcup_n D_n = \bigcup_n A_n$:

Setze $D_1 := A_1$, $D_2 := A_2 \setminus (A_1 \cap A_2) \in \mathscr{D}$, also $D_1 \dot\cup D_2 = A_1 \cup A_2$,

gegeben paarw. disj. $D_1, \dots, D_n \in \mathscr{D}$ mit $D_1 \dot\cup \cdots \dot\cup D_n = A_1 \cup \cdots \cup A_n$ setze

$$D_{n+1} := A_{n+1} \setminus \Big(\dot{\bigcup_{j=1}^{n}} D_j \cap A_{n+1} \Big) \in \mathscr{D},$$

dann ist $\dot\bigcup_{j=1}^{n+1} D_j = \bigcup_{j=1}^{n+1} A_j$. □

Beobachtung und Definition 1.13. $\mathscr{D}_i$, $i \in I$ (beliebige Indexmenge) Dynkin-Systeme (über Ω), so ist $\bigcap_{i\in I} \mathscr{D}_i$ wiederum ein Dynkin-System.

Für $\mathscr{E} \subset 2^\Omega$ heißt

$$\delta(\mathscr{E}) := \bigcap \big\{ \mathscr{D} : \mathscr{D} \text{ ist Dynkin-System mit } \mathscr{E} \subset \mathscr{D} \big\}$$

das von $\mathscr{E}$ erzeugte Dynkin-System. Dies ist offenbar das kleinste Dynkin-System über Ω, das $\mathscr{E}$ umfasst.

Satz 1.14. *$\mathscr{E} \subset 2^\Omega$ $\cap$-stabil, so gilt $\delta(\mathscr{E}) = \sigma(\mathscr{E})$.*

Beweis. Stets gilt: $\delta(\mathscr{E}) \subset \sigma(\mathscr{E})$.

Zeige: $\delta(\mathscr{E})$ ist $\cap$-stabil.

Sei $B \in \delta(\mathscr{E})$,

$$\mathscr{D}_B := \{ A \subset \Omega : A \cap B \in \delta(\mathscr{E}) \}$$

ist ein Dynkin-System:

1) $\Omega \cap B = B \in \delta(\mathscr{E})$, also $\Omega \in \mathscr{D}_B$.

2) Sei $A \in \mathscr{D}_B$. $A^c \cap B = B \setminus (A \cap B) \in \delta(\mathscr{E})$, also $A^c \in \mathscr{D}_B$.

3) Seien $A_1, A_2, A_3, \dots \in \mathscr{D}_B$ paarweise disjunkt, dann ist

$$\Big(\dot{\bigcup_{n=1}^{\infty}} A_n \Big) \cap B = \dot{\bigcup_{n=1}^{\infty}} \underbrace{(A_n \cap B)}_{\in \delta(\mathscr{E})} \in \mathscr{D}_B.$$

Es gilt $\mathscr{E} \subset \mathscr{D}_B$: Betrachte zunächst den Fall $E \in \mathscr{E}$, dann gilt $A \cap E \in \mathscr{E} \subset \delta(\mathscr{E})$ für jedes $A \in \mathscr{E}$ (denn $\mathscr{E}$ ist nach Voraussetzung $\cap$-stabil), also gilt $\mathscr{E} \subset \mathscr{D}_E$ in diesem Fall.

Somit gilt auch $\delta(\mathscr{E}) \subset \mathscr{D}_E$ für jedes $E \in \mathscr{E}$.

Für $B \in \delta(\mathscr{E})$, $E \in \mathscr{E}$ ist daher $E \cap B \in \delta(\mathscr{E})$. Mithin gilt $E \in \mathscr{D}_B$ für $E \in \mathscr{E}, B \in \delta(\mathscr{E})$, d.h. $\mathscr{E} \subset \mathscr{D}_B$.

Da $\mathscr{D}_B$ für jedes $B \in \delta(\mathscr{E})$ selbst ein Dynkin-System ist, folgt $\delta(\mathscr{E}) \subset \mathscr{D}_B$, d.h. $\delta(\mathscr{E})$ ist $\cap$-stabil wie behauptet.

Mit Prop. 1.12 folgt $\delta(\mathscr{E}) = \sigma(\mathscr{E})$. $\square$

Bemerkung 1.15 (Algebren, Ringe, Halbringe)**.** Die Mengenoperationen $A \Delta B := (A \setminus B) \cup (B \setminus A)$ („symmetrische Differenz“) und $A \cap B$ erfüllen die Axiome eines kommutativen Rings im Sinne der Algebra (Δ entspricht „+“, $\cap$ entspricht „×“, $\varnothing$ ist die additive 0, Ω ist die multiplikative 1).

Ein *Ring* $\mathscr{A}$ (im Sinne der Mengensysteme) erfüllt

$$\begin{aligned} &\textit{i)} && \varnothing \in \mathscr{A}, \\ &\textit{ii)} && A, B \in \mathscr{A} \;\Rightarrow\; A \cup B \in \mathscr{A}, \\ &\textit{iii)} && A, B \in \mathscr{A} \;\Rightarrow\; A \setminus B \in \mathscr{A} \end{aligned}$$

insbesondere gilt für einen Ring $\mathscr{A}$ auch

$$A, B \in \mathscr{A} \Rightarrow A \cap B = A \setminus (A \setminus B) \in \mathscr{A}$$

und $\mathscr{A}$ ist abgeschlossen unter endlichen Vereinigungen.

Ein *Halbring* $\mathscr{A}$ (im Sinne der Mengensysteme) erfüllt $\varnothing \in \mathscr{A}$, für $A, B \in \mathscr{A}$ gilt $A \cap B \in \mathscr{A}$ und $A \setminus B = \dot\bigcup_{j=1}^{m} C_j$ für geeignete paarweise disjunkte $C_1, \ldots, C_m \in \mathscr{A}$

Ein „klassisches" Beispiel eines (Mengen-)Rings über $\mathbb{R}$ ist

$$\{\text{endl. Vereinigungen von Intervallen aus } \mathbb{R}\},$$

ein „klassisches" Beispiel eines (Mengen-)Halbrings über $\mathbb{R}$ ist

$$\{\text{endl. Intervalle aus } \mathbb{R}\}.$$

1.2 Mengenfunktionen, Maße

Definition 1.16. $(\Omega, \mathscr{A})$ ein messbarer Raum. Eine Abbildung $\mu : \mathscr{A} \to [0, \infty]$ heißt ein *Maß* (auf $(\Omega, \mathscr{A})$), wenn gilt $\mu(\varnothing) = 0$ und

$$A_1, A_2, \dots \in \mathscr{A} \text{ paarw. disjunkt } \Rightarrow \mu\Big(\dot\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n)$$

(μ ist σ-additiv).

μ heißt *endliches Maß*, wenn $\mu(\Omega) < \infty$; μ heißt *Wahrscheinlichkeitsmaß*, wenn $\mu(\Omega) = 1$; μ heißt *σ-endlich*, wenn es eine Folge $(C_n)_{n \in \mathbb{N}} \subset \mathscr{A}$ gibt mit $\bigcup_n C_n = \Omega$ und $\mu(C_n) < \infty$ für alle n.

Wenn $\mathscr{A}$ (nur) eine Algebra ist: $\mu : \mathscr{A} \to [0, \infty]$ mit

$$\mu\big(A_1 \dot\cup A_2 \dot\cup \cdots \dot\cup A_n\big) = \mu(A_1) + \cdots + \mu(A_n)$$

heißt ein *Inhalt* (d.h. ein Inhalt μ ist endlich additiv); μ heißt in diesem Fall ein *Prämaß*, wenn es σ-additiv ist.

Bemerkung 1.17. Sei μ Maß

1. $A, B \in \mathscr{A}, A \subset B \Rightarrow \mu(A) \le \mu(B)$ (denn $\mu(B) = \mu(A) + \mu(B \setminus A)$)
2. $\mu(A_1 \cup A_2 \cup \cdots \cup A_n) \le \mu(A_1) + \mu(A_2) + \cdots + \mu(A_n)$
3. Wenn $\mu(\Omega) < \infty$, so gilt für $A, B \in \mathscr{A}$

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$$

4. Wenn $\mu(\Omega) < \infty$, so gilt für $A_1, \dots, A_n \in \mathscr{A}, n \in \mathbb{N}$

$$\mu(A_1 \cup \dots \cup A_n) = \sum_{\varnothing \neq K \subset \{1,\dots,n\}} (-1)^{|K|-1} \mu\Big(\bigcap_{k \in K} A_k \Big)$$

(„Einschluss-Ausschluss-Regel").

Beispiel. 1. Für beliebiges Ω und $a \in \Omega$ definiert für $A \in \mathscr{A}$

$$\delta_a(A) = \begin{cases} 1, & a \in A, \\ 0, & a \notin A \end{cases}$$

das *Dirac-Maß* δ_a im Punkt a.

2. Für abzählbares Ω (ausgestattet mit $\mathscr{A} = 2^\Omega$) ist

$$\mu(A) = \#A, \quad A \subset \Omega$$

das *Zählmaß.*

Satz 1.18. *Seien μ_1, μ_2 Maße auf messbarem Raum $(\Omega, \mathscr{A})$, $\mathscr{E}$ ein $\cap$-stabiler Erzeuger von $\mathscr{A}$), es gelte*

$$\mu_1(C) = \mu_2(C) \quad \textit{für alle } C \in \mathscr{E}$$

und es gebe $(C_n)_n \subset \mathscr{E}$ mit $C_1 \subset C_2 \subset \dots$, $\bigcup_n C_n = \Omega$, $\mu_1(C_n) < \infty$.

Dann gilt $\mu_1 = \mu_2$ auf $\mathscr{A}$.

Beweis. Fixiere $E \in \mathscr{E}$ mit $\mu_1(E)\,(= \mu_2(E)\,) < \infty$,

$$\mathscr{D}_E := \big\{ A \in \mathscr{A} : \mu_1(A \cap E) = \mu_2(A \cap E) \big\}$$

ist ein Dynkin-System: $\Omega \in \mathscr{D}_E$ n. Vor.; sei $A \in \mathscr{D}_E$, so ist :

$$\begin{aligned} \mu_1(A^c \cap E) &= \mu_1(E) - \mu_1(A \cap E) \\ &= \mu_2(E) - \mu_2(A \cap E) = \mu_2(A^c \cap E), \end{aligned}$$

d.h. auch $A^c \in \mathscr{D}_E$; seien $A_1, A_2, A_3, \dots \in \mathscr{D}_E$ paarweise disjunkt :

$$\begin{aligned} &\mu_1\Big(\big(\dot{\bigcup_{n=1}^{\infty}} A_n \big) \cap E \Big) = \mu_1\Big(\dot{\bigcup_{n=1}^{\infty}} (A_n \cap E) \Big) = \sum_{n=1}^{\infty} \mu_1(A_n \cap E) \\ &= \sum_{n=1}^{\infty} \mu_2(A_n \cap E) = \mu_2\Big(\big(\dot{\bigcup_{n=1}^{\infty}} A_n \big) \cap E \Big), \end{aligned}$$

d.h. $\dot{\bigcup}_{n=1}^{\infty} A_n \in \mathscr{D}_E$.

N. Vor. ist $\mathscr{E} \subset \mathscr{D}_E$, also (mit Satz 1.14)

$$\mathscr{D}_E \supset \delta(\mathscr{E}) = \sigma(\mathscr{E}) = \mathscr{A}.$$

Sei $(C_n)_n \subset \mathscr{E}$ mit $C_n \nearrow \Omega$ wie oben, $B_n \coloneqq C_n \setminus C_{n-1}$ ($C_0 \coloneqq \varnothing$), also

$$C_N = \dot{\bigcup_{n=1}^{N}} B_n \quad \text{für jedes } N \in \mathbb{N}.$$

Sei $A \in \mathscr{A}$: Für $i = 1, 2$ ist

$$\begin{aligned}
\mu_i(A) &= \mu_i\Big(A \cap \dot{\bigcup_{n=1}^{\infty}} B_n\Big) = \mu_i\Big(\dot{\bigcup_{n=1}^{\infty}}(A \cap B_n)\Big) \\
&= \sum_{n=1}^{\infty} \mu_i(A \cap B_n) = \lim_{N\to\infty} \sum_{n=1}^{N} \mu_i(A \cap B_n) \\
&= \lim_{N\to\infty} \mu_i(A \cap C_N)
\end{aligned}$$

und der Ausdruck in der letzten Zeile hängt nach obigem nicht von i ab.

Somit gilt $\mu_1(A) = \mu_2(A)$ für alle $A \in \mathscr{A}$. □

Beispiel 1.19 (W'maße auf $\mathbb{R}^d$ und Verteilungsfunktionen)**.** μ W'maß auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, $F_\mu : \mathbb{R}^d \to [0, 1]$,

$$F_\mu(x) = \mu\big((-\infty, x]\big), \quad x = (x_1, \ldots, x_d) \in \mathbb{R}^d$$

ist die *Verteilungsfunktion* von μ.

μ ist durch F_μ festgelegt (nach Satz 1.18, denn $\sigma\big((-\infty, x] : x \in \mathbb{R}^d\big) = \mathcal{B}(\mathbb{R}^d)$ und $\{(-\infty, x] : x \in \mathbb{R}^d\}$ ist $\cap$-stabil).

Proposition 1.20. *$\mathscr{G} \subset 2^\Omega$ Algebra, μ endlich additiv auf $\mathscr{G}$ mit $\mu(\varnothing) = 0$. Wir betrachten folgende Eigenschaften:*

- *i)* μ *σ-additiv (auf $\mathscr{G}$)*
- *ii)* μ *aufsteigend stetig*: $(A_n)_n \subset \mathscr{G}$ *mit* $A_n \underset{n\to\infty}{\nearrow} A \in \mathscr{G} \Rightarrow \mu(A_n) \xrightarrow[n\to\infty]{} \mu(A)$
- *iii)* μ *absteigend stetig*: $(A_n)_n \subset \mathscr{G}$ *mit* $A_n \underset{n\to\infty}{\searrow} A \in \mathscr{G}$ *und* $\mu(A_n) < \infty \Rightarrow \mu(A_n) \xrightarrow[n\to\infty]{} \mu(A)$
- *iv)* μ *stetig in* $\varnothing$: $(A_n)_n \subset \mathscr{G}$ *mit* $A_n \underset{n\to\infty}{\searrow} \varnothing$ *und* $\mu(A_n) < \infty \Rightarrow \mu(A_n) \xrightarrow[n\to\infty]{} 0$

Stets gilt i) $\iff$ ii), iii) $\iff$ iv), ii) $\Rightarrow$ iii).

Wenn $\mu(\Omega) < \infty$ so gilt auch iii) $\Rightarrow$ ii), d.h. dann sind i)–iv) äquivalent.

Beweis. „*i)* $\iff$ *ii)*“ : $\mathscr{G} \ni A_n \nearrow A \in \mathscr{G}$, setze $B_n \coloneqq A_n \setminus A_{n-1}$ (mit $A_0 \coloneqq \varnothing$), dann ist $\dot{\bigcup}_{n=1}^N B_n = A_N \nearrow_{N\to\infty} A$.

Wenn *i)* gilt, so ist

$$\mu(A) = \mu\Big(\dot{\bigcup_n} B_n\Big) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{N\to\infty} \sum_{n=1}^{N} \mu(B_n) = \lim_{N\to\infty} \mu(A_N),$$

d.h. dann gilt auch *ii)*.

Analog: Seien $B_1, B_2, \dots \in \mathscr{G}$ paarweise disjunkt mit

$$A_N := B_1 \cup \dots \cup B_N \underset{N\to\infty}{\nearrow} \dot{\bigcup_n} B_n =: A \in \mathscr{G}.$$

Wenn *ii)* gilt, so gilt wie oben $\mu(A) = \mu\Big(\dot\bigcup_n B_n\Big) = \sum_{n=1}^{\infty} \mu(B_n)$, d.h. dann gilt auch *i)*.

„*iii)* $\Rightarrow$ *iv)*“ : ✓

„*iv)* $\Rightarrow$ *iii)*“ : Sei $A_n \searrow A \in \mathscr{G}$, $\mu(A_n) < \infty$. $\mathscr{G} \ni B_n := A_n \setminus A \searrow_{n\to\infty} \varnothing$ und $\mu(B_n) \le \mu(A_n) < \infty$. Somit

$$\mu(A_n) = \mu(A) + \mu(B_n) \underset{n\to\infty}{\longrightarrow} \mu(A),$$

denn $\mu(B_n) \to 0$ n. Vor.

„*ii)* $\Rightarrow$ *iii)*“ : Sei $A_n \searrow A \in \mathscr{G}$, $\mu(A_n) < \infty$, setze $B_n := A_1 \setminus A_n \nearrow_{n\to\infty} A_1 \setminus A$ (und $A_1 = B_n \cup A_n$), somit (nach *ii)*)

$$\mu(A_n) = \mu(A_1) - \mu(B_n) \underset{n\to\infty}{\longrightarrow} \underbrace{\mu(A_1)}_{=\mu(A)+\mu(A_1\setminus A)} - \mu(A_1 \setminus A) = \mu(A),$$

d.h. *iii)* gilt.

Sei weiter $\mu(\Omega) < \infty$, dann gilt „*iii)* $\Rightarrow$ *ii)*“ : $A_n \nearrow A \in \mathscr{G}$, also $B_n := \Omega \setminus A_n \searrow \Omega \setminus A =: B$,

$$\mu(A_n) = \mu(\Omega) - \mu(B_n) \underset{n\to\infty}{\longrightarrow} \mu(\Omega) - \mu(B) = \mu(A).$$

$\square$

Definition 1.21. Ein Tripel $(\Omega, \mathscr{A}, \mu)$, wo $\mathscr{A}$ eine σ-Algebra über Ω ($\neq \varnothing$) und μ ein Maß auf $\mathscr{A}$ ist, heißt ein *Maßraum*. Wenn μ ein Wahrscheinlichkeitsmaß ist, so heißt $(\Omega, \mathscr{A}, \mu)$ (auch) ein *Wahrscheinlichkeitsraum*.

Satz 1.22 (Carathéodorys[2] Maßfortsetzungssatz). *$\mathscr{A} \subset 2^\Omega$ Algebra, μ Prämaß auf $\mathscr{A}$. Dann gibt es ein Maß $\widetilde{\mu}$ auf $\sigma(\mathscr{A})$, das auf $\mathscr{A}$ mit μ übereinstimmt.*

Wenn μ σ-endlich ist, so ist $\widetilde{\mu}$ dadurch eindeutig bestimmt.

Definition 1.23. $\nu : 2^\Omega \to [0, \infty]$ mit $\nu(\varnothing) = 0$,

i) $\forall\, C_1 \subset C_2 \subset \Omega : \quad \nu(C_1) \le \nu(C_2) \quad$ („Monotonie“)

ii) $\forall\, (C_n)_{n\in\mathbb{N}} \subset 2^\Omega : \quad \nu\Big(\bigcup_{n=1}^{\infty} C_n\Big) \le \sum_{n=1}^{\infty} \nu(C_n) \quad$ („σ-Subadditivität“)

heißt ein *äußeres Maß*.

Lemma 1.24. *μ Prämaß auf Algebra $\mathscr{A}$.*

$$\mu^*(C) := \inf\Big\{ \sum_{n=1}^{\infty} \mu(A_n) : (A_n)_{n\in\mathbb{N}} \subset \mathscr{A} \text{ mit } \bigcup_n A_n \supset C \Big\}, \quad C \subset \Omega$$

ist ein äußeres Maß mit $\mu^*(A) = \mu(A)$ für $A \in \mathscr{A}$

[2] Constantin Carathéodory, 1873–1950

Beweis. 1) Zeige $\mu^* = \mu$ auf $\mathscr{A}$: Sei $A \in \mathscr{A}$, wähle $\varepsilon > 0$.

Nach Def. gibt es $(A_n)_{n\in\mathbb{N}} \subset \mathscr{A}$ mit $A \subset \bigcup_n A_n$ und

$$\mu^*(A) \leq \sum_n \mu(A_n) \leq \mu^*(A) + \varepsilon$$

o.E. mit $A_n \subset A$ und $A_1, A_2, \ldots$ paarweise disjunkt (sonst gehe über zu $A_n' := A_n \cap A \in \mathscr{A}$ und dann zu $A_n'' := A_n' \setminus \left(\bigcup_{j=1}^{n-1} A_j'\right) \in \mathscr{A}$), somit $A = \dot\bigcup_{n=1}^{\infty} A_n$ und

$$\mu^*(A) \leq \sum_n \mu(A_n) = \mu(A) \leq \mu^*(A) + \varepsilon,$$

mit $\varepsilon \downarrow 0$ folgt $\mu^*(A) = \mu(A)$.

Insbesondere gilt $\mu^*(\varnothing) = \mu(\varnothing) = 0$.

2) Sei $C \subset C' \subset \Omega$, so gilt $\mu^*(C) \leq \mu^*(C')$, denn $C' \subset \bigcup_n A_n \Rightarrow C \subset \bigcup_n A_n$.

3) Sei $(C_n)_n \subset 2^\Omega$, wähle $\varepsilon > 0$.

Für $n \in \mathbb{N}$ gibt es $A_{n,i} \in \mathscr{A}$, $i \in \mathbb{N}$ mit $C_n \subset \bigcup_{i=1}^{\infty} A_{n,i}$ und

$$\mu^*(C_n) \leq \sum_{i=1}^{\infty} \mu(A_{n,i}) \leq \mu^*(C_n) + \frac{\varepsilon}{2^n}.$$

Sei B_m, $m \in \mathbb{N}$ eine Aufzählung von $\{A_{n,i} : n, i \in \mathbb{N}\}$, es ist

$$\bigcup_n C_n \subset \bigcup_{n,i} A_{n,i} = \bigcup_m B_m,$$

also

$$\begin{aligned}
\mu^*\Big(\bigcup_n C_n\Big) &\leq \sum_m \mu(B_m) = \sum_n \sum_i \mu(A_{n,i}) \\
&\leq \sum_n \Big(\mu^*(C_n) + \frac{\varepsilon}{2^n}\Big) = \varepsilon + \sum_n \mu^*(C_n).
\end{aligned}$$

Mit $\varepsilon \downarrow 0$ folgt σ-Subadditivität von μ^*. □

Definition 1.25. μ^* ein äußeres Maß. $A \subset \Omega$ heißt μ^**-messbar*, wenn gilt

$$\forall\, C \subset \Omega : \ \ \mu^*(C) \geq \mu^*(A \cap C) + \mu^*(A^c \cap C)$$

(wegen σ-Subadditivität von μ^* gilt dann tatsächlich $\mu^*(C) = \mu^*(A \cap C) + \mu^*(A^c \cap C)$ für $C \subset \Omega$).

Beweis von Satz 1.22. Definiere μ^* wie in Lemma 1.24, setze

$$\mathscr{A}_* := \{A \subset \Omega : A \text{ ist } \mu^*\text{-messbar}\}.$$

1) Zeige $\mathscr{A} \subset \mathscr{A}_*$: Sei $A \in \mathscr{A}$ fest. Zu $C \subset \Omega$, $\varepsilon > 0$ sei $(A_n)_n \subset \mathscr{A}$ mit

$$\mu^*(C) \leq \sum_n \mu(A_n) \leq \mu^*(C) + \varepsilon.$$

Es ist

$$A \cap C \subset \bigcup_n (\underbrace{A \cap A_n}_{\in \mathscr{A}}), \quad A^c \cap C \subset \bigcup_n (\underbrace{A^c \cap A_n}_{\in \mathscr{A}}),$$

also

$$\begin{aligned} \mu^*(A \cap C) + \mu^*(A^c \cap C) &\leq \sum_n \mu(A \cap A_n) + \sum_n \mu(A^c \cap A_n) \\ &= \sum_n \mu\big((A \cap A_n) \cup (A^c \cap A_n)\big) = \sum_n \mu(A_n) \leq \mu^*(C) + \varepsilon, \end{aligned}$$

mit $\varepsilon \downarrow 0$ folgt $\mu^*(C) \geq \mu^*(A \cap C) + \mu^*(A^c \cap C)$, d.h. die Behauptung.

2) Zeige: $\mathscr{A}_*$ ist eine Algebra: $\varnothing \in \mathscr{A}_*$ ✓; $A \in \mathscr{A}_* \Rightarrow A^c \in \mathscr{A}_*$ ✓(Def. 1.25 ist symmetrisch bzgl. Vertauschung von A und A^c)

Seien nun $A, B \in \mathscr{A}_*$: Für $C \subset \Omega$ ist

$$\begin{aligned} \mu^*(C) &= \mu^*(B \cap C) + \mu^*(B^c \cap C) \\ &= \mu^*(A \cap B \cap C) + \mu^*(A^c \cap B \cap C) + \mu^*(A \cap B^c \cap C) + \mu^*(A^c \cap B^c \cap C). \end{aligned} \tag{1.1}$$

Ersetzen wir C durch $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ in (1.1), so ergibt sich

$$\mu^*\big((A \cup B) \cap C\big) = \mu^*(A \cap B \cap C) + \mu^*(A^c \cap B \cap C) + \mu^*(A \cap B^c \cap C) \tag{1.2}$$

(denn $\mu^*\big((A^c \cap B^c \cap (A \cup B) \cap C\big) = \mu^*(\varnothing) = 0$). Einsetzen von (1.2) in (1.1) liefert

$$\begin{aligned} \mu^*(C) &= \mu^*\big((A \cup B) \cap C\big) + \mu^*\big(A^c \cap B^c \cap C\big) \\ &= \mu^*\big((A \cup B) \cap C\big) + \mu^*\big((A \cup B)^c \cap C\big), \end{aligned}$$

da dies für beliebiges C stimmt, gilt $A \cup B \in \mathscr{A}_*$.

3) Zeige: $\mathscr{A}_*$ ist ein Dynkin-System.

Seien $A_1, A_2, \dots \in \mathscr{A}$ paarw. disjunkt, $A := \dot\bigcup_{n=1}^\infty A_n$. Für $A_1, A_2 \in \mathscr{A}_*$ mit $A_1 \cap A_2 = \varnothing$ und beliebiges $C \subset \Omega$ liefert (1.2)

$$\mu^*\big((A_1 \mathbin{\dot\cup} A_2) \cap C\big) = \mu^*(A_1^c \cap A_2 \cap C) + \mu^*(A_1 \cap A_2^c \cap C),$$

induktiv folgt für $n \in \mathbb{N}$

$$\mu^*\Big(\big(\dot\bigcup_{i=1}^n A_i\big) \cap C\Big) = \sum_{i=1}^n \mu^*(A_i \cap C) \quad \text{für alle } C \subset \Omega.$$

Für $n \in \mathbb{N}$ ist $\dot\bigcup_{i=1}^n A_i \in \mathscr{A}_*$ nach 2), also gilt für jedes $C \subset \Omega$

$$\begin{aligned} \mu^*(C) &= \mu^*\Big(\big(\dot\bigcup_{i=1}^n A_i\big) \cap C\Big) + \mu^*\Big(\underbrace{\big(\dot\bigcup_{i=1}^n A_i\big)^c \cap C}_{\supset A^c \cap C}\Big) \\ &\geq \sum_{i=1}^n \mu^*(A_i \cap C) + \mu^*(A^c \cap C). \end{aligned}$$

Mit $n \to \infty$ folgt

$$\mu^*(C) \geq \sum_{i=1}^{\infty} \mu^*(A_i \cap C) + \mu^*(A^c \cap C). \tag{1.3}$$

Andererseits ist

$$\begin{aligned}\mu^*(C) = \mu^*\big((A \cap C) \cup (A^c \cap C)\big) &= \mu^*\Big(\big(\dot{\bigcup_{i=1}^{\infty}}(A_i \cap C)\big) \cup (A^c \cap C)\Big)\\ &\leq \sum_{i=1}^{\infty} \mu^*(A_i \cap C) + \mu^*(A^c \cap C) \end{aligned} \tag{1.4}$$

(wir verwenden die σ-Subadditivität von μ^*). (1.3) und (1.4) liefern

$$\mu^*(C) = \sum_{i=1}^{\infty} \mu^*(A_i \cap C) + \mu^*(A^c \cap C), \tag{1.5}$$

da $\mu^*(A \cap C) = \mu^*\Big(\big(\dot\bigcup_{i=1}^{\infty}(A_i \cap C)\big)\Big) \leq \sum_{i=1}^{\infty} \mu^*(A_i \cap C)$ und $\mu^*(C) \leq \mu^*(A \cap C) + \mu^*(A^c \cap C)$ gilt (nochmals σ-Subadditivität von μ^*), impliziert dies

$$\mu^*(C) = \mu^*(A \cap C) + \mu^*(A^c \cap C),$$

also $A \in \mathscr{A}_*$.

4) 2) und 3) zeigen mit Prop. 1.12, dass $\mathscr{A}_*$ eine σ-Algebra ist; wegen 1) gilt $\sigma(\mathscr{A}) \subset \mathscr{A}_*$.

Zeige: μ^* ist σ-additiv auf $\mathscr{A}_*$:

Seien $A_1, A_2, \dots \in \mathscr{A}_*$ paarweise disjunkt, setze $A := \dot\bigcup_{i=1}^{\infty} A_i \in \mathscr{A}_*$ an Stelle von C in (1.5) ein:

$$\mu^*\Big(\dot{\bigcup_{i=1}^{\infty}} A_i\Big) = \sum_{i=1}^{\infty} \mu^*(A_i)$$

(denn $A \cap A_i = A_i$ n. Vor., $\mu^*(A^c \cap A) = \mu^*(\emptyset) = 0$). Somit leistet $\widetilde{\mu} := \mu^*|_{\sigma(\mathscr{A})}$ das Gewünschte.

5) Zur Eindeutigkeit: Seien $A_1 \subset A_2 \subset \dots \in \mathscr{A}$ mit $A_n \nearrow \Omega$, $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$, $\widetilde{\mu}_1, \widetilde{\mu}_2$ zwei Maße auf $\sigma(\mathscr{A})$ mit $\mu = \widetilde{\mu}_1 = \widetilde{\mu}_2$ auf $\mathscr{A}$. Dann gilt $\widetilde{\mu}_1 = \widetilde{\mu}_2$ auf $\sigma(\mathscr{A})$ nach Satz 1.18, da $\mathscr{A}$ $\cap$-stabil ist. □

Erinnerung 1.26. Eine *Verteilungsfunktion* (einer Wahrscheinlichkeitsverteilung auf $\mathbb{R}$) ist eine Funktion $F : \mathbb{R} \to [0,1]$, die nicht-fallend und rechtsstetig ist mit $\lim_{x\to\infty} F(x) = 1$, $\lim_{x\to-\infty} F(x) = 0$.

Satz 1.27. *Zu jeder Verteilungsfunktion F gehört genau ein W'maß μ_F auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit*

$$\mu_F\big((-\infty, x]\big) = F(x), \quad x \in \mathbb{R}.$$

Beweis. Betrachte

$$J_1 := \{(a,b] : a < b\},\ J_2 := \{(-\infty,b] : b \in \mathbb{R}\},\ J_3 := \{(a,\infty) : a \in \mathbb{R}\}, \quad J := J_1 \cup J_2 \cup J_3,$$
$$\mathscr{G} := \Big\{ \dot{\bigcup_{j=1}^{n}} I_j : n \in \mathbb{N}, I_j \in J \Big\} \cup \{\varnothing\}.$$

$\mathscr{G}$ *ist* eine Algebra.

Für $A = \dot\bigcup_{k=1}^{n} I_k \in \mathscr{G}$ mit $I_k = (a_k, b_k]\,(\cap \mathbb{R})$ definieren wir

$$\mu(A) := \sum_{k=1}^{n} \big(F(b_k) - F(a_k)\big)$$

(mit Setzungen $F(-\infty) = 0$, $F(\infty) = 1$).

μ ist wohldefiniert: Wenn man $I_k = (a_k, b_k]$ ersetzt durch $I_k = I_k' \,\dot\cup\, I_k''$ mit $I_k' = (a_k, c_k]$, $I_k'' = (c_k, b_k]$ und einem $c_k \in (a_k, b_k)$, so ändert sich wegen

$$F(b_k) - F(a_k) = \big(F(b_k) - F(c_k)\big) + \big(F(c_k) - F(a_k)\big)$$

der Wert von $\mu_F(A)$ nicht.

Es gilt $\mu(\varnothing) = 0$ und $\mu(\mathbb{R}) = 1$, μ *ist* additiv.

Zur σ-Additivität: Nach Prop. 1.20 genügt es zu zeigen, dass μ stetig in $\varnothing$ ist.
a) (nahezu Kompaktheit)
Zeige: Für $I \in J$ und $\varepsilon > 0$ gibt es $u, v \in \mathbb{R}$, $u \le v$ mit

$$[u,v] \subset I \quad \text{und} \quad \mu\big((u,v]\big) > \mu(I) - \varepsilon$$

Für $I = (a,b] \in J_1$ verwende $v = b$ und verwende, dass wegen Rechtsstetigkeit von F gilt

$$\lim_{u \downarrow a} \mu\big((u,v]\big) = \lim_{u \downarrow a} F(v) - F(u) = F(b) - F(a) = \mu\big((a,b]\big),$$

für $I \in J_2$ bzw. $I \in J_3$ verwende $\lim_{x \to -\infty} F(x) = 0$ bzw. $\lim_{x \to \infty} F(x) = 1$.

Analog: Für allgemeines $A \in \mathscr{G}$ und $\varepsilon > 0$ gibt es $B \in \mathscr{G}$ beschränkt mit $\overline{B} \subset A$ und $\mu(A \smallsetminus B) < \varepsilon$.

b) Seien $A_n \in \mathscr{G}$, $A_1 \supset A_2 \supset \cdots$ mit $\inf_{n \in \mathbb{N}} \mu(A_n) \ge \varepsilon$ für ein $\varepsilon > 0$.
Zeige: Dann gilt $\bigcap_n A_n \neq \varnothing$

Für $n \in \mathbb{N}$ wähle nach a) $B_n \in \mathscr{G}$, B_n beschränkt (also $\overline{B_n}$ kompakt) mit $B_n \subset \overline{B_n} \subset A_n$ und $\mu(A_n \smallsetminus B_n) \le 2^{-n-1}\varepsilon$, es ist

$$\mu\Big(A_n \smallsetminus \bigcap_{k=1}^{n} B_k\Big) \le \mu\Big(\bigcup_{k=1}^{n} (A_k \smallsetminus B_k)\Big) \le \sum_{k=1}^{n} 2^{-k-1}\varepsilon = \frac{\varepsilon}{2} \sum_{k=1}^{n} 2^{-k} < \frac{\varepsilon}{2}$$

(für die erste Ungleichung verwende $A_k \supset A_n$ für $k \le n$), demnach

$$\mu\Big(\bigcap_{k=1}^{n} B_k\Big) = \mu(A_n) - \mu\Big(A_n \smallsetminus \bigcap_{k=1}^{n} B_k\Big) \ge \frac{\varepsilon}{2},$$

somit ist

$$K_n := \bigcap_{k=1}^n \overline{B_k} \supset \bigcap_{k=1}^n B_k \neq \varnothing$$

kompakt und nicht leer mit $K_n \subset A_n$, $K_n \supset K_{n+1}$.

Es folgt

$$\bigcap_{n=1}^\infty A_n \supset \bigcap_{n=1}^\infty K_n \neq \varnothing$$

(Cantor'scher Durchschnittsatz).

□

Lemma 1.28. $(\Omega, \mathscr{A})$ *messbarer Raum,* $\mu_1, \mu_2, \dots$ *Maße darauf,* $c_1, c_2, \dots \geq 0$*, so ist* $\mu := \sum_n c_n \mu_n$ *ebenfalls ein Maß auf* $(\Omega, \mathscr{A})$.

Wenn alle μ_n *W'maße sind und* $\sum_n c_n = 1$*, so ist auch* μ *ein W'maß.*

Beweis. Offenbar ist $\mu(A) = \sum_n c_n \mu_n(A) \geq 0$ für $A \in \mathscr{A}$ und $\mu(\varnothing) = 0$.

Zur σ-Additivität: Seien $A_1, A_2, \dots \in \mathscr{A}$ paarw. disjunkt,

$$\begin{aligned} \mu\Big(\dot{\bigcup_{j=1}^\infty} A_j\Big) &= \sum_n c_n \mu_n\Big(\dot{\bigcup_{j=1}^\infty} A_j\Big) = \sum_n c_n \sum_j \mu_n(A_j) \\ &= \sum_j \sum_n c_n \mu_n(A_j) = \sum_j \mu(A_j) \end{aligned}$$

□

Satz 1.29 (Konstruktion des (Borel-)Lebesgue-Maßes[3] auf $\mathbb{R}$). *Es gibt genau ein Maß* λ *auf* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *mit* $\lambda\big((a,b]\big) = b - a$ *für* $a < b$.

Beweis. 1) $F_1(x) := (x \wedge 1) \vee 0$ ist Verteilungsfunktion, sei $\lambda_{(0,1]}$ das nach Satz 1.27 zugehörige W'maß.

Offenbar gilt $\lambda_{(0,1]}(A) = 0$ wenn $A \subset (0,1]^c$,

$$\lambda_{(0,1]}\big((u,v]\big) = F_1(v) - F_1(u) = v - u \quad \text{für } 0 \leq u < v \leq 1.$$

Analog betrachte für $m \in \mathbb{Z}$: $F_m(x) := \big((x - m + 1) \wedge 1\big) \vee 0$ mit zugehörigem Maß $\lambda_{(m-1,m]}$; es erfüllt

$$\lambda_{(m-1,m]}(A) = \lambda_{(m-1,m]}\big((A \cap (m-1,m]\big), \quad A \in \mathcal{B}(\mathbb{R}^d)$$

und $\lambda_{(m-1,m]}\big((u,v]\big) = (v \wedge m) - (u \vee (m-1))$.

$\lambda := \sum_{m \in \mathbb{Z}} \lambda_{(m-1,m]}$ ist ein Maß mit

$$\lambda\big((a,b]\big) = \sum_{m=\lfloor a \rfloor}^{\lceil b \rceil} \big(b \wedge (m+1)\big) - (a \vee m) = b - a \quad \text{für } a < b.$$

[3]Henri Lebesgue, 1875–1941

Eindeutigkeit folgt aus Satz 1.18 (die Menge der Intervalle $(a,b]$ ist ein $\cap$-stabiler Erzeuger von $\mathcal{B}(\mathbb{R}^d)$, vgl. Bsp. 1.8; λ ist σ-endlich, wähle z.B. $C_n := [-n,n]$, $n \in \mathbb{N}$ als $\mathbb{R}$ ausschöpfende Folge). $\square$

Bemerkung 1.30. 1) Das Lebesgue-Maß λ auf $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ ist translationsinvariant, d.h. $\lambda(A+x) = \lambda(A)$ für $A \in \mathcal{B}(\mathbb{R})$, $x \in \mathbb{R}$.

2) (Analogon von Satz 1.27 für allgemeine endliche Maße) Jede beschränkte, nicht-fallende, rechtsstetige Funktion $F : \mathbb{R} \to [0,\infty)$ mit $\lim_{x\to-\infty} F(x) = 0$ steht via

$$\mu_F\big((-\infty,x]\big) = F(x), \quad x \in \mathbb{R}$$

in eindeutiger Beziehung zu einem endlichen Maß μ_F auf $(\mathbb{R},\mathcal{B}(\mathbb{R}))$.

3) (Lebesgue-Stieltjes-Maße) Zu $F : \mathbb{R} \to \mathbb{R}$ nicht-fallend und rechtsstetig gibt es genau ein Maß μ_F auf $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ mit

$$\mu_F\big((a,b]\big) = F(b) - F(a), \quad -\infty < a < b < \infty.$$

Betrachte dazu (analog zum Beweis von Satz 1.29) $F_m(x) = F\big((x \wedge m) \vee (m-q)\big) - F(m-1)$, $m \in \mathbb{Z}$ und zugehöriges μ_{F_m} aus 2), $\mu_F := \sum_{m\in\mathbb{Z}} \mu_{F_m}$ leistet das Gewünschte.

Eindeutigkeit folgt aus Satz 1.18.

Zum d-dimensionalen Fall:

Definition 1.31. $F : \mathbb{R}^d \to [0,1]$ heißt eine *d-dimensionale Verteilungsfunktion*, wenn es folgenden Bedingungen genügt:

1. F rechtsstetig, d.h. $x_n \searrow x$ (koordinatenweise) $\Rightarrow F(x_n) \to F(x)$
2. $F(x_n) \to 1$ wenn $x_n \to (+\infty,\dots,+\infty)$
3. $F(x_n) \to 0$ wenn $\min_{i=1,\dots,d} x_{n,i} \to -\infty$
4. Für $x = (x_1,\dots,x_d) < y = (y_1,\dots,y_d)$ (koordinatenweise), parametrisiere die Ecken des d-Quaders $(x,y]$ mit $\{1,2\}^d$ via $\big\{u = (z_1^{(i_1)},\dots,z_d^{(i_d)}) : i_1,\dots,i_d \in \{1,2\}\big\}$ wo $z_j^{(1)} = x_j$, $z_j^{(2)} = y_j$, es muss gelten

$$\sum_{u \text{ Ecken}} (-1)^{\#\{1\le m\le d:\, i_m=1\}} F(u) \Big(= \mu_F((x,y]) \Big) \ge 0$$

Satz 1.32. *W'maße auf* $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$ *und d-dimensionale Verteilungsfunktionen F stehen in* 1-1-*Beziehung via*

$$\mu_F\big((-\infty,x]\big) = F(x), \quad x \in \mathbb{R}^d.$$

Beweisskizze. Beweis analog zum Beweis von Satz 1.27, man verwendet als Algebra

$$\mathscr{G} = \Big\{ \dot{\bigcup_{j=1}^{n}} I_{j,1} \times I_{j,2} \times \cdots \times I_{j,d} : n \in \mathbb{N}, I_{j,k} \in J \Big\},$$

die Menge der endlichen Vereinigungen (ggfs. verallgemeinerter) d-dimensionaler Quader.

□

Korollar 1.33 (d-dimensionale Produktmaße)**.** *Seien $\mu_1, \dots, \mu_d$ W'maße auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, dann gibt es genau ein W'maß μ auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ mit*

$$\mu\big((-\infty, x]\big) = \prod_{i=1}^{d} \mu_i\big((-\infty, x_i]\big), \quad x = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Man schreibt $\mu = \mu_1 \otimes \cdots \otimes \mu_d$, es gilt

$$\mu(A_1 \times A_2 \times \cdots \times A_d) = \mu_1(A_1) \cdot \mu_2(A_1) \cdot \cdots \cdot \mu_d(A_d), \quad A_1, \dots, A_d \in \mathcal{B}(\mathbb{R}). \tag{1.6}$$

Beweis. Sei F_i die Verteilungsfunktion von μ_i, $i = 1, \dots, d$,

$$F(x) := F_1(x_1) F_2(x_2) \cdots F_d(x_d), \quad x = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Zeige: F ist d-dim. Verteilungsfunktion:

$a = (a_1, \dots, a_d) < (b_1, \dots b_d) = d$

$$\begin{aligned}\prod_{i=1}^{d} \mathbf{1}(a_i < x_i \le b_i) &= \prod_{i=1}^{d} \big(\mathbf{1}(x_i \le b_i) - \mathbf{1}(x_i \le a_i)\big) \\ &= \sum_{K \subset \{1,\dots,d\}} (-1)^{|K|} \mathbf{1}\big(x \le e(a,b,K)\big)\end{aligned}$$

mit $e(a,b,K)_i = \begin{cases} a_i, & i \in K, \\ b_i, & i \notin K \end{cases}$

F erfüllt

$$\begin{aligned}\sum_{K \subset \{1,\dots,d\}} (-1)^{|K|} F\big(e(a,b,K)\big) &= \sum_{K \subset \{1,\dots,d\}} (-1)^{|K|} \Big(\prod_{i \in K} F_i(a_i)\Big)\Big(\prod_{i \notin K} F_i(b_i)\Big) \\ &= \prod_{i=1}^{d} \big(F(b_i) - F(a_i)\big) \ge 0\end{aligned}$$

für beliebiges $a, b \in \mathbb{R}^d$, $a < b$, d.h. F ist d-dim. Verteilungsfunktion.

Zur Formel (1.6): Nach obigen gilt (1.6), wenn $A_i = (-\infty, x_i]$ für gewisse $x_i \in \mathbb{R}$. Seien $x_2, \dots, x_d$ fixiert, betrachte für $A \in \mathcal{B}(\mathbb{R})$

$$\mu_1'(A) := \mu\big(A \times (-\infty, x_2] \times (-\infty, x_3] \times \cdots \times (-\infty, x_d]\big), \quad \mu_1''(A) := \mu_1(A) \prod_{i=2}^{d} \mu\big((-\infty, x_i]\big).$$

μ_1' und μ_1'' *sind* Maße auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, die auf allen Halbintervallen der Form $(-\infty, x_1]$ übereinstimmen, nach Satz 1.18 gilt somit $\mu_1' = \mu_1''$, d.h. in (1.6) dürfen wir $(-\infty, x_1]$ durch eine allgemeine Borel-Menge A_1 ersetzen.

Fixieren wir nun $A_1 \in \mathcal{B}(\mathbb{R})$, $x_3, \dots, x_d \in \mathbb{R}$ und betrachten für $A \in \mathcal{B}(\mathbb{R})$

$$\mu_2'(A) \coloneqq \mu\big(A_1 \times A \times (-\infty, x_3] \times \cdots \times (-\infty, x_d]\big), \quad \mu_2''(A) \coloneqq \mu_1(A_1)\mu_2(A) \prod_{i=3}^{d} \mu\big((-\infty, x_i]\big),$$

so folgt analog $\mu_2' = \mu_2''$, d.h. in (1.6) dürfen wir $(-\infty, x_1]$ und $(-\infty, x_2]$ jeweils durch eine allgemeine Borel-Menge A_1 bzw. A_2 ersetzen. Iteration dieses Arguments zeigt (1.6) allgemein. □

Satz 1.34 (d-dimensionales Lebesguemaß)**.** *Es gibt genau ein translationsinvariantes Maß λ auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ mit $\lambda\big((0,1]^d\big) = 1$.*

Definition 1.35 (Nullmengen)**.** $(\Omega, \mathscr{A}, \mu)$ Maßraum. $N \in \mathscr{A}$ heißt (μ-)Nullmenge, wenn $\mu(N) = 0$ gilt.

$$\mathscr{N}_\mu \coloneqq \{N \subset \Omega : \exists\, A \in \mathscr{A} \text{ mit } N \subset A \text{ und } \mu(A) = 0\}$$

Offenbar gilt

$$N_1, N_2, \dots \in \mathscr{N}_\mu \;\Rightarrow\; \bigcup_j N_j \in \mathscr{N}_\mu$$

Sprechweisen μ-fast überall (abgekürzt μ-f.ü.), bzw. auch μ-fast sicher, abgekürzt μ-f.s., wenn μ ein W'maß ist: Eine Eigenschaft

$$E(\omega) \text{ gilt } \mu\text{-f.ü., wenn } \big\{\omega \in \Omega : E(\omega) \text{ gilt nicht } \big\} \in \mathscr{N}_\mu$$

Für $A, B \in \mathscr{A}$ schreibt man $A = B \,(\mathrm{mod}\, \mu)$, wenn $A \Delta B \in \mathscr{N}_\mu$.

Bericht 1.36 (Maßvervollständigung, siehe z.B. [Kl, Bem. 1.70])**.**

$$\mathscr{A}_\mu = \sigma\big(\mathscr{A} \cup \mathscr{N}_\mu\big) = \big\{B \subset \Omega : \exists\, A \in \mathscr{A} :\, A \Delta B \in \mathscr{N}_\mu\big\}$$

ist eine σ-Algebra und

$$\mu(B) \coloneqq \mu(A) \quad \text{wenn } A \in \mathscr{A} \text{ mit } A \Delta B \in \mathscr{N}_\mu$$

ist wohldefinierte Fortsetzung von μ auf $\mathscr{A}_\mu$.

1.3 Messbare Funktionen, Zufallsvariablen

Definition 1.37. $(\Omega, \mathscr{A})$, $(\Omega', \mathscr{A}')$ messbare Räume. Eine Abbildung $f : \Omega \to \Omega'$ heißt *$\mathscr{A}$-$\mathscr{A}'$-messbar* (oft auch nur *messbar*, wenn die beteiligten σ-Algebren aus dem Kontext klar sind), wenn gilt

$$\forall\, B \in \mathscr{A}' :\; f^{-1}(B) = \big\{\omega \in \Omega : f(\omega) \in B\big\} \in \mathscr{A}.$$

Beobachtung 1.38. 1 (*Indikatorfunktion*). $A \in \mathscr{A}$, $\mathbf{1}_A : \Omega \to \{0,1\}$,

$$\mathbf{1}_A(\omega) = \begin{cases} 1, & \text{falls } \omega \in A, \\ 0, & \text{falls } \omega \notin A \end{cases}$$

ist messbar ($\mathbf{1}_A^{-1}(\{0\}) = A^c, \mathbf{1}_A^{-1}(\{1\}) = A \in \mathscr{A}$).

2 (Komposition messbarer Abbildungen). $(\Omega_1, \mathscr{A}_1)$, $(\Omega_2, \mathscr{A}_2)$, $(\Omega_3, \mathscr{A}_3)$ messbare Räume, $f : \Omega_1 \to \Omega_2$ und $g : \Omega_2 \to \Omega_3$ seien ($\mathscr{A}_1$-$\mathscr{A}_2$- bzw. $\mathscr{A}_2$-$\mathscr{A}_3$-)messbar. Dann ist $h := g \circ f$ ($\mathscr{A}_1$-$\mathscr{A}_3$-)messbar, denn für $B \in \mathscr{A}_3$ ist

$$h^{-1}(B) = f^{-1}\big(\underbrace{g^{-1}(B)}_{\in \mathscr{A}_2}\big) \in \mathscr{A}_1.$$

Beobachtung und Definition 1.39 (Erzeugte σ-Algebra)**.** Sei $(\Omega', \mathscr{A}')$ messbarer Raum, $\Omega \neq \varnothing$, $h : \Omega \to \Omega'$.

$$\sigma(h) = \left\{h^{-1}(B) : B \in \mathscr{A}'\right\}$$

ist die kleinste σ-Algebra auf Ω, bezüglich der h messbar ist. (Falls $\mathscr{A}$ σ-Algebra auf Ω und h $\mathscr{A}$-$\mathscr{A}'$-messbar, so gilt natürlich $\sigma(h) \subset \mathscr{A}$.)

$\sigma(h)$ heißt die von h *erzeugte σ-Algebra*.

Sind allgemeiner $(\Omega_i, \mathscr{A}_i)$, $i \in I$ messbarere Räume und $h_i : \Omega \to \Omega_i$ (I beliebige Indexmenge), so bezeichnet $\sigma(h_i, i \in I)$ die kleinste σ-Algebra auf Ω, bezüglich der alle h_i, $i \in I$ messbar sind.

Lemma 1.40 (Messbarkeit auf Erzeugermenge genügt)**.** *$f : \Omega \to \Omega'$, $\mathscr{E}' \subset 2^{\Omega'}$ mit $\sigma(\mathscr{E}') = \mathscr{A}'$, so ist f $\mathscr{A}$-$\mathscr{A}'$-messbar g.d.w.*

$$\forall\, B \in \mathscr{E}' : \; f^{-1}(B) = \left\{\omega \in \Omega : f(\omega) \in B\right\} \in \mathscr{A}.$$

Beweis. „$\Rightarrow$“ : ✓
„$\Leftarrow$“ : $\widetilde{\mathscr{A}} := \left\{B \subset \Omega' : f^{-1}(B) \in \mathscr{A}\right\}$ *ist* eine σ-Algebra, n. Vor. gilt $\sigma(\mathscr{E}') = \mathscr{A}'$, also auch $\widetilde{\mathscr{A}} = \mathscr{A}'$. □

Im Fall reellwertiger Funktionen lassen wir oft auch die Werte $+\infty$ oder $-\infty$ zu (Übergang von $\mathbb{R}$ zu $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$). Man spricht dann auch von „numerischen Funktionen“.

Korollar 1.41. *1. Insbesondere für $f : \Omega \to \mathbb{R}$ (wobei $\mathbb{R}$ die Borel-σ-Algebra $\mathcal{B}(\mathbb{R})$ trägt) genügt es (mit Bsp. 1.8) zu prüfen, dass*

$$f^{-1}(I) \in \mathscr{A} \quad \textit{für alle Intervalle } I = [a,b]$$

(oder auch für alle offenen Intervalle, oder für alle Halbintervalle $(-\infty, b]$, $b \in \mathbb{R}$, etc.) gilt.

2 (Topologischer Fall). Wenn $f : \Omega \to E$, wobei E topologischer Raum mit Borel-σ-Algebra $\mathcal{B}(E)$, so genügt für Messbarkeit

$$f^{-1}(\mathcal{O}) \in \mathscr{A} \text{ für jedes offene } \mathcal{O} \subset E.$$

Insbesondere: falls Ω selbst ein topolgischer Raum (versehen mit seiner Borel-σ-Algebra) ist, so sind sind alle stetigen Abbildungen messbar.

Definition 1.42 (Produktraum, Produkt-σ-Algebra)**.** I endl. oder abzählbar, $(S_i, \mathscr{A}_i)$, $i \in I$ messbare Räume $S = \times_{i\in I} S_i = \{s = (s_i)_{i\in I} : s_i \in S_i\}$,

$$\mathscr{A}_\otimes := \sigma\Big(\Big\{ \bigtimes_{i\in I} A_i : A_i \in \mathscr{A}_i \Big\}\Big)$$

(die von den „verallgemeinerten Quadern" erzeugte σ-Algebra) heißt die *Produkt-σ-Algebra* (der $\mathscr{A}_i$).

Mit den Koordinatenabbildungen (oder Projektionsabbildungen) $\pi : S \to S_i$, $\pi_i\big((s_j)_{j\in I}\big) = s_i$ ist $\mathscr{A}_\otimes = \sigma(\pi_i, i \in I)$.

Lemma 1.43 (Messbarkeit im Produktfall)**.** *S wie in Def. 1.42, $f : \Omega \to S$, $f(\omega) = \big(f_i(\omega)\big)_{i\in I}$ ist messbar g.d.w*

$$f_i : \Omega \to S_i \quad \text{messbar für jedes } i \in I.$$

Beweis. „$\Rightarrow$" : $\pi_i : S \to S_i$ *ist* messbar, mit Beob. 1.38 daher auch $f_i = \pi_i \circ f$

„$\Leftarrow$" : Mengen $A = \bigtimes_{i\in I} A_i$ mit $A_i \in \mathscr{A}_i$ erzeugen $\mathscr{A}_\otimes$ und für ein solches A ist

$$f^{-1}(A) = \{\omega \in \Omega : f_i(\omega) \in A_i, i \in I\} = \bigcap_{i\in I} f_i^{-1}(A_i) \in \mathscr{A}_\otimes$$

n. Vor., Beh. folgt mit Lemma 1.40. □

Beobachtung 1.44. $S = \overline{\mathbb{R}}^\infty$, inf : $S \to \mathbb{R}$ mit $\inf\big((x_n)_{n\in\mathbb{N}}\big) = \inf_n x_n$ und sup : $S \to \mathbb{R}$ mit $\sup\big((x_n)_{n\in\mathbb{N}}\big) = \sup_n x_n$ sind $\mathcal{B}(\overline{\mathbb{R}}^\infty)$-$\mathcal{B}(\overline{\mathbb{R}})$-messbar:

Für $x \in \mathbb{R}$ ist

$$\inf{}^{-1}\big([x,\infty]\big) = [x,\infty] \times [x,\infty] \times [x,\infty] \times \cdots \in \mathcal{B}(\overline{\mathbb{R}}^\infty),$$
$$\sup{}^{-1}\big([-\infty,x]\big) = [-\infty,x] \times [-\infty,x] \times [-\infty,x] \times \cdots \in \mathcal{B}(\overline{\mathbb{R}}^\infty).$$

Lemma 1.45. *E topologischer Raum mit Borel-σ-Algebra $\mathcal{B}(E)$, $f_1, f_2, \cdots : \Omega \to E$ seien ($\mathscr{A}$-$\mathcal{B}(E)$-)messbar und*

$$f(\omega) := \lim_{n\to\infty} f_n(\omega) \quad \text{existiere für alle } \omega \in \Omega.$$

Dann ist f ($\mathscr{A}$-$\mathcal{B}(E)$-)messbar.

Beweis. Sei $\mathcal{O} \subset E$ offen, dann ist

$$f^{-1}(\mathcal{O}) = \bigcup_n \bigcap_{m \geq n} f_m^{-1}(\mathcal{O}) \in \mathscr{A}.$$

□

Definition 1.46 (Elementare Funktionen)**.** $(\Omega, \mathscr{A})$ messbarer Raum. $f : \Omega \to \mathbb{R}$ der Form $f(\omega) = \sum_{i=1}^n a_i \mathbf{1}_{A_i}(\omega)$ mit $n \in \mathbb{N}$, $a_1, \dots, a_n \in \mathbb{R}$, $A_1, \dots, A_n \in \mathscr{A}$ heißt eine *elementare Funktion.*

Lemma 1.47. *$(\Omega, \mathscr{A})$ messbarer Raum.*

1. Elementare Funktionen sind messbar.

2. Für $\mathscr{A}$-$\mathcal{B}(\overline{\mathbb{R}})$-messbares $f : \Omega \to [0, \infty]$ gibt es eine Folge von elementaren Funktionen f_n mit $f_n \nearrow f$.

3. Jedes $\mathscr{A}$-$\mathcal{B}(\overline{\mathbb{R}})$-messbare $f : \Omega \to \overline{\mathbb{R}}$ kann punktweise durch elementare Funktionen approximiert werden.

Beweis. 1. Schreibe $f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$ und o.E. seien die A_i paarweise disjunkt mit $\bigcup_{i=1}^n A_i = \Omega$, dann ist $f^{-1}(B) = \bigcup_{i : a_i \in B} A_i \in \mathscr{A}$ für $B \subset \mathbb{R}$.

2. Wähle z.B. $f_n(\omega) = \left(2^{-n} \lfloor 2^n f(\omega) \rfloor\right) \wedge n$.

3. Zerlege $f = f^+ - f^-$, dann verwende 2. □

Lemma 1.48 (Faktorisierungslemma)**.** *$(\Omega, \mathscr{A})$, $(\Omega', \mathscr{A}')$ messbare Räume, $f : \Omega \to \Omega'$ ($\mathscr{A}$-$\mathscr{A}'$-)messbar. Dann ist $g : \Omega \to \overline{\mathbb{R}}$ $\sigma(f)$-$\mathcal{B}(\overline{\mathbb{R}})$-messbar g.d.w.*

es gibt ein $\mathscr{A}'$-$\mathcal{B}(\overline{\mathbb{R}})$ messbares $h : \Omega' \to \overline{\mathbb{R}}$ mit $g = h \circ f$.

Beweis. „$\Leftarrow$“ : ✓

„$\Rightarrow$“ : Sei zunächst $g \geq 0$. Zeige: Es gibt $A_1, A_2, \dots \in \sigma(f)$, $\alpha_1, \alpha_2, \dots \in [0, \infty)$ mit

$$g = \sum_{n=1}^{\infty} \alpha_n \mathbf{1}_{A_n} \tag{1.7}$$

Sei dazu $g_n := \left(2^{-n} \lfloor 2^n g \rfloor\right) \wedge n$, $B_{n,i} := \{\omega : g_n(\omega) - g_{n-1}(\omega) = i 2^{-n}\}$, $i = 0, 1, \dots, 2^n$ ($B_{n,i} \in \sigma(f)$ n. Vor.), also

$$g_n - g_{n-1} = \sum_{i=0}^{2^n} i 2^{-n} \mathbf{1}_{B_{n,i}}$$

Umsummieren $(n, i) \mapsto m \in \mathbb{N}$ liefert

$$g = g_0 + \sum_{n=1}^{\infty} (g_n - g_{n-1}) = \sum_{m=1}^{\infty} \alpha_m \mathbf{1}_{A_m}$$

mit geeigneten $\alpha_m \in [0, \infty)$, $A_m \in \sigma(f)$.

Wegen $A_m \in \sigma(f)$ gibt es $B_m \in \mathscr{A}'$ mit $f^{-1}(B_m) = A_m$, setze

$$h := \sum_{m=1}^{\infty} \alpha_m \mathbf{1}_{B_m}$$

Im allgemeinen Fall wähle zu g^+, g^- wie oben h^+, h^-, setze $h := h^+ - h^-$. □

Definition 1.49 (Bildmaß)**.** $(\Omega, \mathscr{A}, \mu)$ Maßraum, $(\Omega', \mathscr{A}')$ messbarer Raum, $f : \Omega \to \Omega'$ messbar.

$$\mathscr{A}' \ni A' \mapsto \mu\big(f^{-1}(A')\big)$$

ist ein Maß auf $(\Omega', \mathscr{A}')$, es heißt das *Bildmaß* von μ unter f und wird mit $\mu \circ f^{-1}$ bezeichnet (manchmal schreibt man auch $f(\mu)$).

1.3.1 Erinnerung: Wahrscheinlichkeitstheoretische Interpretation

Definition 1.50. $(\Omega, \mathscr{F}, \mathbb{P})$ W'raum, $(S, \mathscr{A})$ messbarer Raum. $A \in \mathscr{F}$ heißen *Ereignisse*, eine ($\mathscr{F}$-$\mathscr{A}$-)messbare Abbildung $X : \Omega \to S$ heißt eine *Zufallsvariable* (mit Werten in S), oft kürzt man Zufallsvariable als ZV ab, auch die Benennung *Zufallsgröße* ist verbreitet.

Für $B \in \mathscr{A}$ schreibt man

$$\{X \in B\} := X^{-1}(B) \quad (\in \mathscr{F})$$

für das Ereignis „X nimmt einen Wert in B an“.

Das Bildmaß $\mathbb{P} \circ X^{-1}$ heißt die *Verteilung* von X, man schreibt auch $\mathscr{L}(X)$ und $X \sim \mu$, falls $\mathbb{P} \circ X^{-1} = \mu$.

Beispiel 1.51. 1) μ W'maß auf $(S, \mathscr{A})$, so ist $X = \mathrm{Id}_S$ ZV auf $(S, \mathscr{A}, \mu)$ mit $X \sim \mu$.

Insbesondere ist $U := \mathrm{Id}_{[0,1]}$ eine uniform auf $[0,1]$ verteilte ZV auf $(\Omega, \mathscr{F}, \mathbb{P}) = ([0,1], \mathcal{B}([0,1]), \lambda_{[0,1]})$.

2) F Verteilungsfunktion auf $\mathbb{R}$, so gibt es eine ZV X auf $([0,1], \mathcal{B}([0,1]), \mathbb{P})$, $\mathbb{P} = \lambda_{[0,1]})$ mit $\mathbb{P}(X \leq x) = F(x)$, $x \in \mathbb{R}$.

Kapitel 2

Unabhängigkeit

Definition 2.1. $(\Omega, \mathscr{F}, \mathbb{P})$ W'raum.

1. $\mathscr{C}_i \subset \mathscr{F}$, $i \in I$ (I beliebige Indexmenge) heißen *unabhängig*, wenn für jedes endliche $J \subset I$, $J \neq \varnothing$

$$\text{für beliebige } A_j \in \mathscr{C}_j,\ j \in J \text{ gilt } \mathbb{P}\Big(\bigcap_{j \in J} A_j\Big) = \prod_{j \in J} \mathbb{P}(A_j).$$

2. Ereignisse $A_i \in \mathscr{F}$, $i \in I$ heißen unabhängig, wenn dies für $\mathscr{C}_i := \{\varnothing, A_i, A_i^c, \Omega\}$ gilt.

3. ZVn X_i, $i \in I$ heißen unabhängig, wenn $\sigma(X_i)$, $i \in I$ unabhängig sind.

Lemma 2.2. *$\mathscr{C}_i \subset \mathscr{F}$, $i \in I$, $\mathscr{C}_i \cup \{\varnothing\}$ seien $\cap$-stabil. Dann gilt*

$$(\mathscr{C}_i, i \in I) \text{ u.a.} \quad \text{g.d.w.} \quad (\sigma(\mathscr{C}_i), i \in I) \text{ u.a.}$$

Beweis. (Es ist nur „$\Rightarrow$" zu zeigen). Wir zeigen: Für $J \subset J' \subset I$, $|J'| < \infty$ gilt

$$\mathbb{P}\Big(\bigcap_{j \in J'} A_j\Big) = \prod_{j \in J'} \mathbb{P}(A_j) \quad \text{für jede Wahl } \begin{cases} A_j \in \sigma(\mathscr{C}_j), & j \in J, \\ A_j \in \mathscr{C}_j, & j \in J' \setminus J \end{cases} \tag{2.1}$$

per Induktion nach $n := |J|$.

Nach Voraussetzung gilt (2.1) für $n = 0$.

Nehmen wir an, (2.1) ist für alle J mit $n := |J|$ erfüllt. Betrachte $\widetilde{J} := J \cup \{j_0\}$ wo $|J| = n$ und $j_0 \in I \setminus J$; sei $J \subset J' \subset I$.

Seien $A_j \in \sigma(\mathscr{C}_j)$ für $j \in J$, $A_j \in \mathscr{C}_j$ für $j \in J' \setminus \widetilde{J}$, betrachte folgende Maße (auf $(\Omega, \mathscr{F})$):

$$\mu(A) := \mu\Big(A \cap \bigcap_{j \in J' \setminus \{j_0\}} A_j\Big), \quad \widetilde{\mu}(A) := \mathbb{P}(A) \cdot \prod_{j \in J' \setminus \{j_0\}} \mathbb{P}(A_j), \quad A \in \mathscr{F}.$$

Nach Induktionsannahme stimmen μ und $\widetilde{\mu}$ auf $\mathscr{C}_{j_0} \cup \{\varnothing, \Omega\}$, das ein $\cap$-stabiler Erzeuger von $\sigma(\mathscr{C}_{j_0})$ ist, überein, mit Satz 1.18 also

$$\mu = \widetilde{\mu} \text{ auf } \sigma(\mathscr{C}_{j_0}),$$

d.h. (2.1) gilt für $n + 1$. □

Bemerkung 2.3. 1. $\mathscr{F} \ni A_i$, $i \in I$ sind u.a. gd.w.

$$\mathbb{P}\Big(\bigcap_{j\in J} A_j\Big) = \prod_{j\in J} \mathbb{P}(A_j) \quad \text{für jedes endliche } J \subset I.$$

2. $\mathscr{C}_i \subset \mathscr{F}$, $i \in I$ u.a., $I = \dot{\bigcup}_{k\in K} I_k$ (wobei K eine beliebige Menge ist), so sind auch $\widetilde{\mathscr{C}_k} := \bigcup_{i\in I_k} \mathscr{C}_i$, $k \in K$ unabhängig.

Beweis. 1. folgt aus Lemma 2.2, denn $\{A_i\}$ ist ein $\cap$-stabiler Erzeuger der σ-Algebra $\{\varnothing, A_i, A_i^c, \Omega\}$.

2. ✓ (zu $A_j \in \widetilde{\mathscr{C}_{k_j}}$, $j = 1, \ldots, n$ mit $k_1, \ldots, k_n$ paarw. verschieden gibt es $i_1, \ldots i_n$ paarw. verschieden mit $A_j \in \mathscr{C}_{i_j}$)

□

Beispiel. 1) $X_1, X_2, \ldots, X_d$ reellwertige ZVn sind u.a. g.d.w.

$$\mathbb{P}\Big(\bigcap_{j=1}^{d} \{X_j \le x_j\}\Big) = \prod_{j=1}^{d} \mathbb{P}(X_j \le x_j) \quad \text{für alle } x = (x_1, \ldots, x_d) \in \mathbb{R}^d$$

(d.h. g.d.w. ihre gemeinsame Verteilungsfunktion Produktgestalt hat).

2) X_i, $i \in I$ diskrete ZVn, d.h. $X_i : \Omega \to S_i$ messbar (wobei S_i endlich oder abzählbar, mit σ-Algebra $\mathscr{A} = 2^{S_i}$).

$(X_i, i \in I)$ unabhängig g.d.w.

$$\text{für } J \subset I, |J| < \infty,\ x_j \in S_j \text{ gilt } \mathbb{P}\Big(\bigcap_{j\in J} \{X_j = x_j\}\Big) = \prod_{j\in J} \mathbb{P}(X_j = x_j).$$

2.0.1 Asymptotische Ereignisse

In diesem Abschnitt geht es um Ereignisse, die gewissermaßen (wenn man bei der Nummerierung an einen Zeitablauf denkt) „unendlich spät“ entschieden werden.

Definition 2.4. Seien $A_1, A_2, \ldots$ Ereignisse,

$$\limsup_{n\to\infty} A_n := \bigcap_{n\ge 1} \bigcup_{m\ge n} A_m$$

(„unendlich viele der A_n treten ein“),

$$\liminf_{n\to\infty} A_n := \bigcup_{n\ge 1} \bigcap_{m\ge n} A_m$$

(„von einem (möglicherweise zufälligen) Index ab treten alle A_n ein“).

Beachte: Es ist $\left(\liminf_{n\to\infty} A_n\right)^c = \limsup_{n\to\infty} A_n^c$, also mit Indikatorvariablen ausgedrückt

$$\limsup_{n\to\infty} \mathbf{1}_{A_n}(\omega) = \mathbf{1}_{\limsup_n A_n}(\omega), \quad \liminf_{n\to\infty} \mathbf{1}_{A_n}(\omega) = \mathbf{1}_{\liminf_n A_n}(\omega).$$

Lemma 2.5 (Lemma von Borel-Cantelli[1]). *Seien $A_1, A_2, \dots$ Ereignisse, $A = \limsup_{n\to\infty} A_n$.*

1. $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty \quad \Longrightarrow \quad P(A) = 0$

2. $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$ *und die $A_1, A_2, \dots$ seien unabhängig* $\quad \Longrightarrow \quad P(A) = 1$

Beweis. 1. Für jedes n ist

$$0 \le \mathbb{P}(A) \le \mathbb{P}\Big(\bigcup_{m\ge n} A_m\Big) \le \sum_{m=n}^{\infty} \mathbb{P}(A_m) \xrightarrow[n\to\infty]{} 0,$$

also gilt $\mathbb{P}(A) = 0$.

2. Es ist

$$\begin{aligned}
\mathbb{P}\Big(\bigcap_{m\ge n} A_m^c\Big) &= \lim_{n'\to\infty} \mathbb{P}\Big(\bigcap_{m=n}^{n'} A_m^c\Big) = \lim_{n'\to\infty} \prod_{m=n}^{n'} \big(1 - \mathbb{P}(A_m)\big) \\
&= \prod_{m=n}^{\infty} \big(1 - \mathbb{P}(A_m)\big) = \exp\Big(\sum_{m=n}^{\infty} \log\big(1 - \mathbb{P}(A_m)\big)\Big) \\
&\le \exp\Big(-\sum_{m=n}^{\infty} \mathbb{P}(A_m)\Big) = 0
\end{aligned}$$

für jedes n (verwende $\log(1-x) \le -x$ für $x \in [0,1]$), also $\mathbb{P}(A^c) \le \sum_{n=1}^{\infty} \mathbb{P}\Big(\bigcap_{m\ge n} A_m^c\Big) = 0$. □

Beispiel 2.6. Seien $X_1, X_2, \dots$ u.a, $\{0,1\}$-wertig mit $\mathbb{P}(X_i = 1) = p_i \in (0,1)$, so gilt

$$\{X_i = 1 \ \infty\text{-oft}\} \text{ f.s.} \quad \text{g.d.w.} \quad \sum_{i=1}^{\infty} p_i = \infty.$$

(Dies ist insbesondere erfüllt, wenn $p_i \equiv p > 0$: wenn man eine p-Münze unendlich oft wirft, wird man mit Sicherheit unendlich viele Erfolge beobachten.

Andererseits: Mit $A_n := \{X_1 = 1\}$, $n \in \mathbb{N}$ ist $\limsup_n A_n = A_1$, also $\mathbb{P}(\limsup_n A_n) = p_1 \in (0,1)$, d.h. man kann in Lemma 2.5, 2. nicht (ohne weitere Voraussetzungen) auf die Unabhängigkeit verzichten.

Definition 2.7. $(\Omega, \mathscr{A})$ messbarer Raum, $\mathscr{A}_1, \mathscr{A}_2, \dots \subset \mathscr{A}$ Teil-σ-Algebren.

$$\mathscr{T} := \cap_{n\ge 1} \mathscr{T}_n \quad \text{mit } \mathscr{T}_n := \sigma(\mathscr{A}_n, \mathscr{A}_{n+1}, \mathscr{A}_{n+2}, \dots)$$

heißt die (von der Folge $\mathscr{A}_1, \mathscr{A}_2, \dots$ erzeugte) *terminale σ-Algebra.*

[1] nach Émile Borel (1871–1956) und Francesco Cantelli (1875–1966)

Beispiel 2.8. X_n reellwertige ZVn (auf $(\Omega, \mathscr{A}, \mathbb{P})$), $\mathscr{A}_n = \sigma(X_n)$. Dann sind

$$\{\limsup_n X_n \le c\}, \{(X_n) \text{ konvergiert}\} = \{\liminf_n X_n \ge \limsup_n X_n\},$$
$$\Big\{\sum_n |X_n| < \infty\Big\}, \Big\{\sum_n X_n \text{ konvergiert}\Big\} \in \mathscr{T}.$$

Satz 2.9 (Kolmogorovs 0-1-Gesetz)**.** *$(\Omega, \mathscr{A}, \mathbb{P})$ W'raum, $\mathscr{A}_1, \mathscr{A}_2, \dots \subset \mathscr{A}$ (unter $\mathbb{P}$) unabhängige Teil-σ-Algebren. Dann gilt*

$$\mathbb{P}(A) \in \{0, 1\} \quad \textit{für alle } A \in \mathscr{T}.$$

(Man sagt: „$\mathscr{T}$ ist trivial“).

Beweis. Sei $\sigma_n := \sigma(\mathscr{A}_1, \dots, \mathscr{A}_n)$, für $n \in \mathbb{N}$ sind σ_n und $\mathscr{T}_{n+1}$ unabhängig (nach Bem. 2.3), also sind auch σ_n und $\mathscr{T} \subset \mathscr{T}_{n+1}$ unabhängig.

$\sigma_\infty := \bigcup_{n=1}^\infty \sigma_n$ ist eine $\cap$-stabile Erzeugermenge von $\sigma(\mathscr{A}_k, k \in \mathbb{N})$ [beachte: für $A, B \in \sigma_\infty$ gibt es ein n mit $A, B \in \sigma_n$, daher auch $A \cap B \in \sigma_n \subset \sigma_\infty$], und σ_∞ ist unabhängig von $\mathscr{T}$, nach Lemma 2.2 also

$$\mathscr{T} \text{ und } \sigma(\mathscr{A}_k, k \in \mathbb{N}) \supset \mathscr{T} \text{ sind u.a.},$$

insbesondere für $A \in \mathscr{T}$:

$$\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A) \cdot \mathbb{P}(A).$$

□

Beispiel. Seien $X_1, X_2, \dots$ u.a. reelle ZVn, dann ist $\mathbb{P}\big(\sum_{n=1}^\infty X_n \text{ konvergiert}\big) \in \{0, 1\}$.

Kapitel 3

Integral und Erwartungswert

Sei $(S, \mathscr{A}, \mu)$ Maßraum.

Notation. Für $f, g : S \to \overline{\mathbb{R}}$ messbar, $A \in \mathcal{B}(\overline{\mathbb{R}})$, $z \in \overline{\mathbb{R}}$ schreiben wir

$$\begin{aligned}
&\{f \in A\} := f^{-1}(A), \quad \{f = z\} := f^{-1}(\{z\}),\\
&\{f = g\} := \{x \in S : f(x) = g(x)\} = (f - g)^{-1}(\{0\}),\\
&\{f \geq g\} := \{x \in S : f(x) \geq g(x)\} = (f - g)^{-1}([0, \infty]), \qquad \text{etc.}
\end{aligned}$$

Definition und Beobachtung 3.1 (Elementares Integral)**.** Für

$$f \in \mathcal{E}^+ := \{\text{nicht-negative elementare Funktionen}\}$$

mit Darstellung $f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$ $(a_i \in \mathbb{R}_+,\ A_i \in \mathscr{A})$ ist

$$I(f) := \sum_{i=1}^n a_i \mu(A_i) = \sum_{z \in f(S)} z \mu\big(\{f = z\}\big)$$

(mit Konvention $0 \cdot \infty = 0$) wohldefiniert.

Beweis. Sei

$$\{C_1, \dots, C_m\} = \Big\{ \bigcap_{j=1}^n B_j : B_j = A_j \text{ oder } B_j = A_j^c \Big\} \setminus \{\varnothing\},$$

d.h. die C_k sind paarweise disjunkt mit $S = \dot\bigcup_{k=1}^m C_k$ und für $i = 1, \dots, n$ gibt es $J_i \subset \{1, \dots, m\}$ mit

$$A_i = \dot\bigcup_{k \in J_i} C_k$$

Für $x \in C_k$ ist

$$f(x) = \sum_{i : i \in J_k} a_i =: \gamma_k$$

Sei $f(S) = \{\gamma_k : k = 1, \dots, m\} = \{z_1, \dots, z_\ell\}$, für $j = 1, \dots, \ell$ setze

$$G_j := \big\{ k \in \{1, \dots, m\} : \gamma_k = z_j \big\}$$

$$\begin{aligned}\sum_i a_i\mu(A_i) &= \sum_i \sum_{k\in J_i} a_i\mu(C_k) = \sum_k \mu(C_k) \sum_{i\,:\,k\in J_i} a_i \\ &= \sum_k \gamma_k \mu(C_k) = \sum_j z_j \mu\Big(\bigcup_{k\in G_j} C_k\Big) = \sum_j z_j \mu\big(\{f = z_j\}\big)\end{aligned}$$

□

Definition 3.2. Für $f : S \to \overline{\mathbb{R}}_+$ heißt

$$\int f\,d\mu := \sup\big\{I(h) : h \in \mathcal{E}^+, h \le f\big\}$$

das Integral[1] von f bezüglich μ.

(Manchmal schreibt man auch $\mu(f) := \int f\,d\mu$ oder auch $\int f(x)\,\mu(dx)$, wenn die „Integrationsvariable" betont werden soll.)

Beobachtung 3.3 (Markov-Ungleichung[2])**.** Für $f : S \to \overline{\mathbb{R}}_+$, $a > 0$ gilt

$$\mu\big(\{f \ge a\}\big) \le \frac{1}{a}\int f\,d\mu$$

(denn $\mathcal{E}^+ \ni h := a\mathbf{1}_{f^{-1}([a,\infty])} \le f$, also $I(h) = a\mu\big(\{f \ge a\}\big) \le \int f\,d\mu$).

Lemma 3.4. *$f, g : S \to \overline{\mathbb{R}}_+$ messbar.*

i) $f \le g$ μ-f.ü. $\implies \int f\,d\mu \le \int g\,d\mu$.

ii) $f = g$ μ-f.ü. $\implies \int f\,d\mu = \int g\,d\mu$.

iii) $f = 0$ μ-f.ü. $\iff \int f\,d\mu = 0$.

iv) $\int f\,d\mu < \infty \implies f < \infty$ μ-f.ü.

Beweis. i) Sei h Elementarfunktion mit $h \le f$, dann ist $\widetilde{h} := h\mathbf{1}_{\{f\le g\}}$ ebenfalls Elementarfunktion und erfüllt $\widetilde{h} \le g$,

$$\begin{aligned} I(\widetilde{h}) &= \sum_{z\,(\in h(S)\smallsetminus\{0\})} z\mu\big(\{\widetilde{h} = z\}\big) = \sum_z \mu\big(\{\widetilde{h} = z\} \cap \{f \le g\}\big) \\ &= \sum_z \mu\big(\{h = z\}\big) = I(h)\end{aligned}$$

die Beh. folgt damit aus Def. 3.2.

ii) Dies folgt aus *i)*.

iii) „⇒" folgt aus *ii)*, zu „⇐": Sei $\int f\,d\mu = 0$, dann ist

$$\begin{aligned} &\mu\big(\{f \ge \tfrac{1}{n}\}\big) \le n\int f\,d\mu = 0 \quad \text{für jedes } n \in \mathbb{N} \text{ (mit Beob. 3.3)}, \\ &\mu\big(\{f > 0\}\big) = \mu\big(\{f \ge 1\}\big) + \sum_{n=1}^{\infty} \mu\big(\{\tfrac{1}{n+1} \le f < \tfrac{1}{n}\}\big) = 0\end{aligned}$$

[1] Präziser: das Lebesgue-Integral, nach Henri Lebesgue (1875–1941)
[2] nach Andrei Andrejewich Markov (1856–1922)

iv) $h := n\mathbf{1}_{\{f=\infty\}} \le f$ für jedes $n \in \mathbb{N}$, also

$$I(h) = n\mu\big(\{f = \infty\}\big) \le \int f\,d\mu < \infty$$

und

$$\mu\big(\{f = \infty\}\big) \le \frac{1}{n}\int f\,d\mu \underset{n\to\infty}{\longrightarrow} 0.$$

□

Lemma 3.5. *f_n, $n \in \mathbb{N}$ messbar, $0 \le f_1 \le f_2 \le \cdots \nearrow f = \sup_{n\in\mathbb{N}} f_n$, so gilt*

$$\int f\,d\mu = \lim_{n\to\infty}\int f_n\,d\mu$$

Beweis. Nach Lemma 3.4 ist $n \mapsto \int f_n\,d\mu$ monoton wachsend und

$$\lim_{n\to\infty}\int f_n\,d\mu \le \int f\,d\mu.$$

Sei $\varepsilon > 0$, h Elementarfunktion mit $0 \le h \le f$,

$$h_n := (h-\varepsilon)^+\mathbf{1}_{\{f_n > f-\varepsilon\}}$$

ist Elementarfunktion mit $0 \le h_n \le f_n$,

$$I(h_n) = \sum_{\tilde{z}} \tilde{z}\mu\big(\{h_n = \tilde{z}\}\big) = \sum_z (z-\varepsilon)^+\mu\big(\{h = z, f_n > f-\varepsilon\}\big) \le \int f_n\,d\mu$$

nach Voraussetzung ist $\{f_n > f-\varepsilon\} \nearrow \{f < \infty\}$ für $n \to \infty$, also

$$\mu\big(\{h = z, f_n > f-\varepsilon\}\big) \underset{n\to\infty}{\nearrow} \mu\big(\{h = z, f < \infty\}\big)$$

somit

$$\sum_z (z-\varepsilon)^+\mu\big(\{h = z, f < \infty\}\big) \le \lim_{n\to\infty}\int f_n\,d\mu$$

für jedes $\varepsilon > 0$, mit $\varepsilon \downarrow 0$ auch

$$I(h\mathbf{1}_{\{f<\infty\}}) = \sum_z z\mu\big(\{h = z, f < \infty\}\big) \le \lim_{n\to\infty}\int f_n\,d\mu$$

Wenn nun $\mu(\{f = \infty\}) = 0$ gilt, so ist $I(h\mathbf{1}_{\{f<\infty\}}) = I(h)$ und Def. 3.2 impliziert $\lim_n \int f_n\,d\mu \ge \int f\,d\mu$, d.h. die Behauptung.

Wenn andererseits $\mu(\{f = \infty\}) > 0$ gilt, so ist $\int f\,d\mu = \infty$ und für jedes $M \in [0,\infty)$ ist

$$\lim_n \int f_n\,d\mu \ge \lim_n \int (f_n \wedge M)\,d\mu = \int (f \wedge M)\,d\mu \ge M\mu(\{f = \infty\}),$$

mit $M \to \infty$ folgt $\lim_n \int f_n\,d\mu = \infty$, d.h. die Behauptung gilt auch in diesem Fall. □

(Gegen-)Beispiel. $(S, \mathscr{A}, \mu) = ([0,1], \mathcal{B}([0,1]), \lambda_{[0,1]})$, $f_n = n\mathbf{1}_{(0,1/n)}$, so gilt

$$\lim_{n\to\infty} f_n(x) = 0 =: f(x) \quad \text{für jedes } x \in [0,1],$$

aber $\displaystyle\int f_n \, d\lambda_{[0,1]} = n\frac{1}{n} = 1 \not\to 0 = \int 0 \, d\lambda_{[0,1]}$

Lemma 3.6. *Sei $f : S \to \overline{\mathbb{R}}_+$ messbar und $f(S)$ höchstens abzählbar, dann gilt*

$$\int f \, d\mu = \sum_{z\in f(S)} z\mu\big(\{f = z\}\big)$$

(mit beliebiger Summationsreihenfolge).

Beweis. Betrachte zunächst f Elementarfunktion: Sei $0 \le h \le f$ Elementarfunktion, dann ist für $z > y$

$$\mu\big(\{f = y, h = z\}\big) = \mu(\varnothing) = 0,$$

also

$$\begin{aligned} I(h) &= \sum_z z\mu\big(\{h = z\}\big) = \sum_z \sum_y z\mu\big(\{f = y, h = z\}\big) \\ &\le \sum_z \sum_y y\mu\big(\{f = y, h = z\}\big) = \sum_y y\mu\big(\{f = y\}\big) \ \big(= I(f)\big) \end{aligned}$$

d.h. die Formel gilt in diesem Fall.

Allgemeiner Fall: Sei $y_1, y_2, \dots$ irgendeine Aufzählung von $f(S)$, $(z_n)_n \subset \overline{\mathbb{R}}_+ \setminus f(S)$ mit $z_n \nearrow \infty$,

$$f_n := \sum_{i=1}^n y_i \mathbf{1}_{\{f=y_i\}} + z_n \mathbf{1}_{\{f=\infty\}}$$

also $0 \le f_1 \le f_2 \le \dots \nearrow f$, Beh. folgt mit Lemma 3.5. □

Lemma 3.7. *$f, g : S \to \overline{\mathbb{R}}_+$ m.b., $\alpha, \beta \in \overline{\mathbb{R}}_+$, so gilt*

$$\int \alpha f + \beta g \, d\mu = \alpha \int f \, d\mu + \beta \int g \, d\mu.$$

Beweis. 1) Seien zunächst $f(S), g(S)$ abzählbar (und somit auch $(f+g)(S)$ abzählbar):

$$\begin{aligned} \int \alpha f + \beta g \, d\mu &= \sum_z z\mu\big(\{\alpha f + \beta g = z\}\big) \\ &= \sum_z z \sum_{\substack{u,v: \\ \alpha u + \beta v = z}} \mu\big(\{f = u, g = v\}\big) = \sum_{u,v} (\alpha u + \beta v)\mu\big(\{f = u, g = v\}\big) \\ &= \alpha \sum_{u,v} u\mu\big(\{f = u, g = v\}\big) + \beta \sum_{u,v} v\mu\big(\{f = u, g = v\}\big) \\ &= \alpha \sum_u u\mu\big(\{f = u\}\big) + \beta \sum_v v\mu\big(\{g = v\}\big) = \alpha \int f \, d\mu + \beta \int g \, d\mu \end{aligned}$$

nach Lemma 3.6.

2) Allgemeiner Fall: $f_n := 2^{-n}\lfloor 2^n f \rfloor \wedge n$, $g_n := 2^{-n}\lfloor 2^n g \rfloor \wedge n$ nehmen jeweils nur endlich viele Werte an, $f_n \nearrow f, g_n \nearrow g$ für $n \to \infty$. Beh. folgt mit 1) und Lemma 3.5. □

Definition 3.8. $f : S \to \overline{\mathbb{R}}$ messbar heißt *μ-integrierbar*, wenn $\int |f|\, d\mu < \infty$. Man setzt dann

$$\int f\, d\mu := \int f^+\, d\mu - \int f^-\, d\mu \quad (\in \mathbb{R})$$

$$\mathcal{L}^1(\mu) := \{f : f \ \mu\text{-integrierbar}\}$$

Für $f \notin \mathcal{L}^1(\mu)$ mit $\int f^+\, d\mu \wedge \int f^-\, d\mu < \infty$ setzt man ebenso, wobei dann die Werte $+\infty$ oder $-\infty$ vorkommen (beachte: ist wohldefiniert).

Man schreibt auch

$$\int f(x)\, \mu(dx) := \mu(f) := \int f\, d\mu,$$

abkürzend oft auch für $A \in \mathscr{A}$

$$\int_A f\, d\mu := \int f \mathbf{1}_A\, d\mu.$$

Satz 3.9. *$f, g \in \mathcal{L}^1(\mu)$*

i) $|f| < \infty$ *μ-f.ü.*, $\int f\, d\mu = 0 \iff f = 0$ *μ-f.ü.*

ii) $f \leq g$ *μ-f.ü.* $\implies \int f\, d\mu \leq \int g\, d\mu$, *insbesondere:* $f = g$ *μ-f.ü.* $\implies \int f\, d\mu = \int g\, d\mu$

iii) $\left| \int f\, d\mu \right| \leq \int |f|\, d\mu$

iv) $\int af + bg\, d\mu = a \int f\, d\mu + b \int g\, d\mu$ *für* $a, b \in \mathbb{R}$

Beweis. i) Dies folgt aus Lemma 3.4, *iii).*

ii) $f \leq g$ f.ü. $\iff$ $f^+ \leq g^+$ f.ü. und $f^- \geq g^-$ f.ü., Beh. folgt aus Lemma 3.4, *i).*

iii) $\left| \int f\, d\mu \right| = \left| \int f^+\, d\mu - \int f^-\, d\mu \right| \leq \left| \int f^+\, d\mu \right| + \left| \int f^-\, d\mu \right| = \int \underbrace{f^+ + f^-}_{=|f|}\, d\mu$

iv) Folgt aus Lemma 3.7 (zerlege $f = f^+ - f^-$, $g = g^+ - g^-$). □

Bemerkung 3.10. 1. $\|f\|_1 := \int |f|\, d\mu$ definiert somit eine Halbnorm auf $\mathcal{L}^1(\mu)$, definiert man $f \sim g :\Leftrightarrow f = g$ μ-f.ü., so ist $L^1(\mu) := \mathcal{L}^1(\mu)/\sim$ ein normierter Raum.

2. Im diskreten Fall $S = \{x_1, x_2, \dots\}$ (mit $\mathscr{A} = 2^S$), $\mu = \sum_{n=1}^\infty a_n \delta_{x_n}$ mit $a_n \geq 0$ ist

$$\mathcal{L}^1(\mu) = \left\{ f : S \to \overline{\mathbb{R}} : \sum_{n=1}^\infty a_n |f(x_n)| < \infty \right\}$$

Satz 3.11 (Monotone Konvergenz, Satz von (Beppo) Levi[3])**.** *$f_1, f_2, \dots \in \mathcal{L}^1(\mu)$, $f_n \nearrow f$ μ-f.ü., dann gilt*

$$\lim_{n\to\infty} \int f_n\, d\mu = \int f\, d\mu$$

(die Gleichung ist möglicherweise als $+\infty = +\infty$ zu lesen).

[3] Beppo Levi, 1875–1961

Beweis. Sei $N \in \mathscr{A}$ mit $\mu(N) = 0$ und

$$f_n(x) \underset{n\to\infty}{\nearrow} f(x) \text{ für alle } x \in S \setminus N.$$

$$0 \le \widetilde{f_n} := \mathbf{1}_{N^c}(f_n - f_1) \underset{n\to\infty}{\nearrow} \mathbf{1}_{N^c}(f - f_1)$$

also (mit Lemma 3.5)

$$\underbrace{\int \widetilde{f_n}\, d\mu}_{=\int f_n\, d\mu - \int f_1\, d\mu} \quad d\mu \underset{n\to\infty}{\longrightarrow} \int \mathbf{1}_{N^c}(f - f_1)\, d\mu = \underbrace{\int f - f_1\, d\mu}_{=\int f\, d\mu - \int f_1\, d\mu}$$

□

Satz 3.12 (Lemma von Fatou[4]). *$f, f_1, f_2, \dots \in \mathcal{L}^1(\mu)$ mit $f_n \ge f$ μ-f.ü., so gilt*

$$\int \liminf_n f_n\, d\mu \le \liminf_n \int f_n\, d\mu$$

Beweis. O.E. seien $f_1, f_2, \dots \ge 0$ f.ü., sonst gehe über zu $f_n - f$.

Sei $g_n := \inf_{m\ge n} f_m$, dann gilt $\int g_n\, d\mu \le \int f_n\, d\mu$ für jedes n und $g_n \underset{n\to\infty}{\nearrow} \liminf_{k\to\infty} f_k$, also

$$\int \liminf_{k\to\infty} f_k\, d\mu = \lim_{n\to\infty} \int g_n\, d\mu \le \liminf_n \int f_n\, d\mu$$

(wir verwenden Satz 3.11 für das Gleichheitszeichen). □

Satz 3.13 (Dominierte Konvergenz, Konvergenzsatz von Lebesgue). *$f, f_1, f_2, \dots$ messbare Funktionen, es gelte*

$$f_n \to f \ \mu\text{-f.ü.} \quad \textit{und} \quad |f_n| \le g \ \mu\text{-f.ü. für ein } g \in \mathcal{L}^1(\mu).$$

Dann gilt $f, f_1, f_2, \dots \in \mathcal{L}^1(\mu)$ und

$$\int \big|f_n - f\big|\, d\mu \underset{n\to\infty}{\longrightarrow} 0, \qquad \textit{insbesondere} \int f_n\, d\mu \underset{n\to\infty}{\longrightarrow} \int f\, d\mu.$$

Beweis. $\int |f_n|\, d\mu, \int |f|\, d\mu \le \int g\, d\mu < \infty$, d.h. $f, f_1, f_2, \dots \in \mathcal{L}^1(\mu)$.

$$0 \le 2g - |f - f_n| \underset{n\to\infty}{\longrightarrow} 2g \quad \mu\text{-f.ü.},$$

also

$$\int 2g\, d\mu \le \liminf_{n\to\infty} \int 2g - |f - f_n|\, d\mu = \underbrace{\int 2g\, d\mu}_{\in[0,\infty)} - \limsup_{n\to\infty} \underbrace{\int |f - f_n|\, d\mu}_{\ge 0}$$

somit $\limsup_{n\to\infty} \int |f - f_n|\, d\mu = 0$.

Schließlich

$$\lim_{n\to\infty} \left| \int f_n\, d\mu - \int f\, d\mu \right| \le \lim_{n\to\infty} \int |f - f_n|\, d\mu = 0$$

□

[4] Pierre Fatou, 1878–1929

Satz 3.14 (Maßtransformation und Integral)**.** *$(S,\mathscr{A}),(S',\mathscr{A}')$ messbare Räume, μ Maß auf $(S,\mathscr{A})$, $f : S \to S'$ messbar, $\mu' := \mu \circ f^{-1}$*

Für $g \in \mathcal{L}^1(\mu')$ oder $g : S' \to \overline{\mathbb{R}}_+$ gilt

$$\int g\, d\mu' = \int g \circ f\, d\mu \tag{3.1}$$

Beobachtung 3.15 (Monotonieprinzip)**.** $(S,\mathscr{A})$ messbarer Raum, sei $\mathscr{K} \subset \{f : f : S \to \overline{\mathbb{R}}_+ \text{ messbar}\}$ mit

i) $a, b \in \mathbb{R}_+,\ f, g \in \mathscr{K} \implies af + bg \in \mathscr{K}$

ii) $(f_n)_{n\in\mathbb{N}} \subset \mathscr{K},\ f_n \nearrow f \implies f \in \mathscr{K}$

iii) $\mathbf{1}_A \in \mathscr{K}$ für $A \in \mathscr{A}$

Dann gilt $\mathscr{K} = \{f : f : S \to \overline{\mathbb{R}}_+ \text{ messbar}\}$.

(*Beweis*: Approximiere allgemeines messbares f wachsend via $\mathscr{K} \ni f_n = 2^{-n}\lfloor 2^n f\rfloor \wedge n \nearrow f$.)

Beweis von Satz 3.14. Sei $\mathscr{K}_+ = \{g : g : S' \to \overline{\mathbb{R}}_+ \text{ messbar, (3.1) gilt für } g\}$.

Für $g = 1_{A'}$ mit $A' \in \mathscr{A}'$ gilt

$$\int g\, d\mu' = \mu'(A') = \mu\big(f^{-1}(A')\big) = \int g \circ f\, d\mu$$

nach Def., also $1_{A'} \in \mathscr{K}_+$.

$\mathscr{K}_+$ erfüllt Eigenschaft *i)* aus Beob 3.15 wegen der Linearität des Integrals (vgl. Satz 3.9) und Eigenschaft *ii)* wegen des Satzes von der monotonen Konvergenz (Satz 3.11). Somit

$$\mathscr{K}_+ = \{g : g : S' \to \overline{\mathbb{R}}_+ \text{ messbar}\}$$

Für $g \in \mathcal{L}^1(\mu')$ gilt (3.1) jeweils für g^+ und g^-. $\square$

Bemerkung und Schreibweise. Für $(\Omega, \mathscr{F}, \mathbb{P})$ W'raum, $X \in \mathcal{L}^1(\mathbb{P})$ oder $X : \Omega \to \overline{\mathbb{R}}_+$ messbar schreibt man

$$\mathbb{E}[X] := \int X\, d\mathbb{P} \quad \text{„Erwartungswert" von } X$$

Für eine $(S',\mathscr{A}')$-wertige ZV Y mit Verteilung ν, $g : S' \to \overline{\mathbb{R}}$, $g \in \mathcal{L}^1(\nu)$ gilt

$$\mathbb{E}\big[g(Y)\big] = \int g \circ Y\, d\mathbb{P} = \int g\, d\nu$$

Definition 3.16. μ, ν Maße auf $(S,\mathscr{A})$, $h : S \to \overline{\mathbb{R}}_+$ messbar heißt (eine) *Dichte* von ν bezüglich μ, wenn gilt

$$\nu(A) = \int \mathbf{1}_A h\, d\mu, \quad A \in \mathscr{A}.$$

Man schreibt $\nu = h\mu$, symbolisch auch $d\nu = h d\mu$, $\dfrac{d\nu}{d\mu} = h$.

Lemma 3.17. *μ Maß auf $(S,\mathscr{A})$, $h : S \to \overline{\mathbb{R}}_+$ messbar, so definiert*

$$\nu(A) := \int \mathbf{1}_A h\, d\mu, \quad A \in \mathscr{A}$$

ein Maß auf $(S,\mathscr{A})$ (ν ist endlich, wenn $h \in \mathcal{L}^1(\mu)$), für $f : S \to \overline{\mathbb{R}}$ m.b. mit $f \geq 0$ μ-f.ü. oder $fh \in \mathcal{L}^1(\mu)$ gilt

$$\int f\, d\nu = \int fh\, d\mu \tag{3.2}$$

Beweis. Offenbar $\nu(\varnothing) = 0$, seien $A_1, A_2, \dots \in \mathscr{A}$ paarw. disjunkt:

$$\nu\Big(\dot{\bigcup_n} A_n\Big) = \int \mathbf{1}_{\dot\bigcup_n A_n} h\, d\mu = \int \sum_n \mathbf{1}_{A_n} h\, d\mu = \sum_n \int \mathbf{1}_{A_n} h\, d\mu = \sum_n \nu(A_n)$$

mit Satz von der monotonen Konvergenz (Satz 3.11).

$$\{f \,:\, f : S \to \overline{\mathbb{R}}_+ \text{ messbar, (3.2) gilt für } f\}$$

erfüllt die Voraussetzungen des Monotonieprinzips (Beob 3.15), also gilt (3.2) für jedes messbare $f \geq 0$.

Falls $fh \in \mathcal{L}^1(\mu)$, so gilt (3.2) mit $hf^+ = (hf)^+$ und $hf^- = (hf)^-$ separat. □

Lemma 3.18. *$\nu = h\mu = h'\mu$ mit $h, h' : S \to \overline{\mathbb{R}}_+$ m.b., ν σ-endlich $\Longrightarrow h = h'$ μ-f.ü.*

Beweis. Sei ν endlich.

$$\begin{aligned}
&\nu\big(\{h > h'\}\big) + \int (h - h')^+ \mu \\
&= \int \mathbf{1}_{\{h>h'\}} h'\, d\mu + \int \mathbf{1}_{\{h>h'\}} (h - h')\, d\mu \\
&= \int \mathbf{1}_{\{h>h'\}} h\, d\mu = \nu\big(\{h > h'\}\big),
\end{aligned}$$

also $(h - h')^+ = 0$ μ-f.ü., analog $(h - h')^- = 0$ μ-f.ü.

Im allgemeinen Fall seien $\mathscr{A} \ni A_n \nearrow S$ mit $\nu(A_n) < \infty$, obiges zeigt

$$\int \mathbf{1}_{A_n} (h - h')^+\, d\mu = \int \mathbf{1}_{A_n} (h - h')^-\, d\mu = 0$$

für jedes n. □

Definition 3.19. Für $1 \leq p < \infty$ sei

$$\mathcal{L}^p(\mu) := \Big\{f : S \to \overline{\mathbb{R}} \text{ messbar} : |f|^p \text{ ist } \mu\text{-integrierbar}\Big\}$$

und für $f \in \mathcal{L}^p(\mu)$ sei $\|f\|_p := \Big(\int |f|^p\, d\mu\Big)^{1/p}$.

$$\|f\|_\infty := \inf\big\{M \in \mathbb{R}_+ : |f| \leq M \ \mu\text{-f.ü.}\big\}$$

heißt *essentielles Supremum* (von f bezüglich μ), $\mathcal{L}^\infty(\mu) := \{f : S \to \overline{\mathbb{R}} \text{ m.b.} : \|f\|_\infty < \infty\}$.

Satz 3.20 (Hölder-Ungleichung[5]). *$p, q \in [1, \infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$, $f \in \mathcal{L}^p(\mu), g \in \mathcal{L}^q(\mu)$*

Dann ist $f \cdot g \in \mathcal{L}^1(\mu)$ mit

$$\|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q.$$

(Für $p = q = 2$ ist dies die Cauchy-Schwarz-Ungleichung.)

Beweis. Die Fälle $p = 1, q = \infty$ (oder umgekehrt) sowie $\|f\|_p = 0$ oder $\|g\|_q = 0$ sind trivial, seien also

$$p, q > 1, \quad \alpha := \|f\|_p, \, \beta := \|g\|_q > 0$$

Für $a, b > 0$ gilt

$$\log(ab) = \frac{1}{p}\log(a^p) + \frac{1}{q}\log(b^q) \leq \log\Big(\frac{1}{p}a^p + \frac{1}{q}b^q\Big)$$

(dies folgt aus der Konkavität der Logarithmusfunktion), also

$$\frac{|f(x)|}{\alpha} \cdot \frac{|g(x)|}{\beta} \leq \frac{1}{p}\frac{|f(x)|^p}{\alpha^p} + \frac{1}{q}\frac{|g(x)|^q}{\beta^q},$$

somit

$$\frac{1}{\alpha\beta}\int |fg| \, d\mu \leq \frac{1}{p} + \frac{1}{q} = 1.$$

□

Satz 3.21 (Jensen'sche Ungleichung[6]). *$(S, \mathscr{A}, \mu)$ W'raum, $k : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ konvex, $f \in \mathcal{L}^1(\mu)$. Dann ist $\int k \circ f \, d\mu$ wohldefiniert und es gilt*

$$\int k \circ f \, d\mu \geq k\Big(\int f \, d\mu\Big)$$

Beweis. Erinnerung:

$$k \text{ ist konvex} \quad \Longleftrightarrow \quad \forall\, x, y \in \mathbb{R}, \alpha \in [0,1] : \; \alpha k(x) + (1-\alpha)k(y) \geq k(\alpha x + (1-\alpha)y)$$

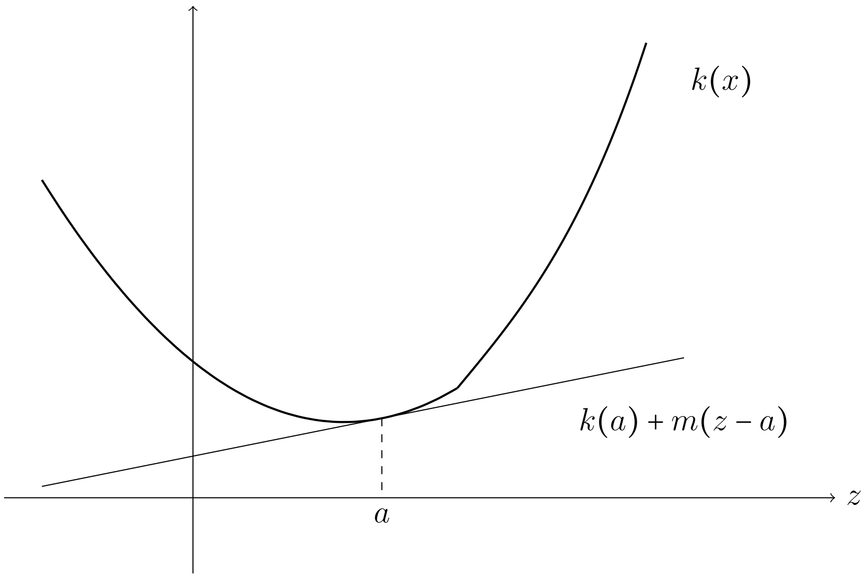
und zu jedem $a \in \mathbb{R}$ gibt es $m = m(a) \in \mathbb{R}$, so dass

$$\forall\, z \in \mathbb{R} : \; k(a) + m(z-a) \leq k(z)$$

gilt (eine konvexe Funktion liegt oberhalb jeder ihrer Tangenten) .

[5] Otto Hölder, 1859–1937

[6] nach Johan Ludvig Jensen, 1859–1925 benannt



Seien a, m wie oben, so ist

$$\big(k \circ f\big)^- \leq \big(k(a) + m(f-a)\big)^- \leq |k(a)| + |m|\big(|f| + |a|\big)$$

also $\int \big(k \circ f\big)^- \, d\mu < \infty$ und

$$\int \big(k \circ f\big)(x)\, \mu(dx) \geq k(a) + m\Big(\int f(x)\, \mu(dx) - a\Big),$$

die Wahl $a = \int f \, d\mu$ liefert die Behauptung.

(Da $f \in \mathcal{L}^1(\mu)$, gilt $\mu(\{|f| < \infty\}) = 1$ und man kann den Wert von $k \circ f$ auf der μ-Nullmenge $\{|f| = \infty\}$ beispielsweise $= 0$ setzen, ohne die Aussage zu ändern. Alternativ kann man beobachten, dass für eine konvexe Funktion k stets $k(\infty) = \lim_{x\to\infty} k(x), k(-\infty) = \lim_{x\to-\infty} k(x) \in \overline{\mathbb{R}}$ existieren.) □

Definition 3.22 (Konvergenzarten)**.** $(S, \mathscr{A}, \mu)$ Maßraum, $f, f_1, f_2, \dots : S \to \overline{\mathbb{R}}$ m.b.

i) $f_n \to f$ μ-fast überall (abgekürzt μ-f.ü.; wenn $\mu(S) = 1$, so sagt man auch „fast sicher“, abgekürzt f.s.), wenn

$$\{x : f_n(x) \not\to f(x)\} = \bigcup_{\varepsilon \in \mathbb{Q} \cap (0,\infty)} \bigcap_n \bigcup_{m \geq n} \big\{|f_m - f| > \varepsilon\big\}$$

eine μ-Nullmenge ist.

ii) $f_n \to f$ im Maß μ $:\Leftrightarrow$ $\forall\, \varepsilon > 0 :\ \mu\big(\{|f_n - f| > \varepsilon\}\big) \underset{n\to\infty}{\longrightarrow} 0$

(man sagt auch „μ-stochastisch“ und schreibt $f_n \to f$ (μ-)stoch.)

iii) $1 \leq p \leq \infty$, $f, f_1, f_2, \dots \in \mathcal{L}^p(\mu)$. $f_n \to f$ im p-ten Mittel (auch „in $\mathcal{L}^p(\mu)$“, geschrieben $f_n \overset{\mathcal{L}^p(\mu)}{\longrightarrow} f$) $:\Leftrightarrow$ $\|f_n - f\|_p \underset{n\to\infty}{\longrightarrow} 0$

Bemerkung 3.23. *i)* $\mu(S) < \infty$, $f_n \to f$ μ-f.ü. $\implies$ $f_n \to f$ im Maß

ii) $f_n \overset{\mathcal{L}^p(\mu)}{\longrightarrow} f$ für ein $p \in [1, \infty]$ $\implies$ $f_n \to f$ im Maß

iii) Die Limiten sind jeweils μ-f.ü. eindeutig bestimmt.

Beweis. *i)* $A_{n,\varepsilon} := \left\{\sup_{m\ge n}|f_m - f| > \varepsilon\right\} \supset A_{n+1,\varepsilon}$, somit

$$\infty > \mu(A_{n,\varepsilon}) \underset{n\to\infty}{\searrow} \mu\Big(\bigcap_n A_{n,\varepsilon}\Big) = 0 \quad \text{für jedes } \varepsilon > 0$$

und $\mu\big(\{|f_n - f| > \varepsilon\}\big) \le \mu(A_{n,\varepsilon})$.

ii) Sei $1 \le p < \infty$:

$$\mu\big(\{|f_n - f| > \varepsilon\}\big) = \mu\big(\{|f_n - f|^p > \varepsilon^p\}\big) \le \frac{1}{\varepsilon^p}\int |f_n - f|^p \, d\mu \underset{n\to\infty}{\longrightarrow} 0$$

mit Markov-Ungleichung (Beob. 3.3).

Für $p = \infty$ beachte, dass $\mu\big(\{|f_n - f| > \varepsilon\}\big) = 0$ für $\varepsilon > \|f_n - f\|_\infty$ gilt.

iii) Es gelte $f_n \to f$ und $f_n \to g$ jeweils im Maß, so ist für $\varepsilon > 0$

$$\mu\big(\{|f - g| > \varepsilon\}\big) \le \mu\big(\{|f_n - f| > \tfrac{\varepsilon}{2}\}\big) + \mu\big(\{|f_n - g| > \tfrac{\varepsilon}{2}\}\big) \underset{n\to\infty}{\longrightarrow} 0.$$

□

(Gegen-)Beispiel. Für ein Gegenbeispiel zur Umkehrung seien $X_1, X_2, \dots$ u.a., $X_n \sim \mathrm{Ber}_{1/n}$, dann gilt

$$X_n \underset{n\to\infty}{\overset{\text{stoch.}}{\longrightarrow}} 0 \qquad \text{(denn für jedes } 1 > \varepsilon > 0 \text{ ist } P\big(|X_n - 0| > \varepsilon\big) = \frac{1}{n} \to 0\text{)},$$

aber

$$P\Big(\underbrace{\{X_n = 1 \ \infty\text{-oft}\}}_{=\limsup_{n\to\infty}\{X_n=1\}}\Big) = 1$$

Lemma 3.24. *$f_1, f_2, \dots : S \to \overline{\mathbb{R}}$ m.b. mit*

$$\lim_{m,n\to\infty} \mu\big(\{|f_m - f_n| > \varepsilon\}\big) = 0 \quad \textit{für alle } \varepsilon > 0.$$

Dann gibt es $n_k \underset{k\to\infty}{\nearrow} \infty$, $f : S \to \overline{\mathbb{R}}$ m.b. mit $f_{n_k} \to f$ μ-f.ü. für $k \to \infty$.

Insbesondere: $f_n \underset{n\to\infty}{\longrightarrow} g$ im Maß μ, so gibt es eine Teilfolge mit $f_{n_k} \underset{k\to\infty}{\longrightarrow} g$ μ-f.ü.

Beweis. Wähle $n_1 < n_2 < \cdots$, so dass für $m \ge n_k$ gilt

$$\mu\big(\{|f_m - f_{n_k}| > 2^{-k}\}\big) \le 2^{-k},$$

insbesondere für $A_k := \left\{|f_{n_{k+1}} - f_{n_k}| > 2^{-k}\right\}$ gilt $\mu(A_k) \le 2^{-k}$ und $\sum_{k=1}^\infty \mu(A_k) < \infty$, mit Borel-Cantelli (Lemma 2.5) also

$$\mu\Big(\limsup_{k\to\infty} A_k\Big) = 0.$$

Auf $\left(\limsup_{k\to\infty} A_k\right)^c$ gilt

$$\sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| < \infty,$$

d.h. $f_{n_m} = f_{n_1} + \sum_{k=1}^{m-1} \left(f_{n_{k+1}} - f_{n_k}\right)$ konvergiert μ-f.ü.

Zum Zusatz: Beachte

$$\mu\big(\{|f_m - f_n| > \varepsilon\}\big) \le \mu\big(\{|f_m - g| > \tfrac{\varepsilon}{2}\}\big) + \mu\big(\{|f_n - g| > \tfrac{\varepsilon}{2}\}\big) \underset{m,n\to\infty}{\longrightarrow} 0.$$

□

Satz 3.25. $1 \le p \le \infty$, $f_1, f_2, \dots$ *Cauchy-Folge in* $\mathcal{L}^p(\mu)$, *d.h.* $\lim_{m,n\to\infty} \|f_n - f_m\|_p = 0$.

Dann gibt es ein $f \in \mathcal{L}^p(\mu)$ *mit* $f_n \to f$ *in* $\mathcal{L}^p(\mu)$.

Beweis. Betrachte $1 \le p < \infty$: Nach Bem. 3.23, *ii)* und Lemma 3.24 gibt es $f : S \to \overline{\mathbb{R}}$ m.b. und $n_k \nearrow \infty$ mit $f_{n_k} \to f$ μ-f.ü. Es ist

$$\begin{aligned} \int |f_m - f|^p \, d\mu &\le \liminf_{k\to\infty} \int |f_m - f_{n_k}|^p \, d\mu \\ &\le \sup_{n\ge m} \int |f_m - f_n|^p \, d\mu \underset{m\to\infty}{\longrightarrow} 0 \end{aligned}$$

n. Vor. Insbesondere $f = f_m - (f_m - f) \in \mathcal{L}^p(\mu)$ und $\|f_m - f\|_p \to 0$.

□

Satz 3.26 (Minkowski-Ungleichung[7]). $1 \le p \le \infty$, $f, g \in \mathcal{L}^p(\mu)$, *so gilt*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p$$

Beweis. Fälle $p = 1$, $p = \infty$ oder falls $\int |f + g|^p \, d\mu = 0$: ✓

Sei $1 < p < \infty$, $0 < \int |f + g|^p \, d\mu$ $(< \infty)$, setze

$$q := \frac{p}{p-1} \ (> 1), \quad \text{somit } \frac{1}{p} + \frac{1}{q} = 1.$$

$$\begin{aligned} \int |f + g|^p \, d\mu &\le \int |f|\,|f + g|^{p-1} \, d\mu + \int |g|\,|f + g|^{p-1} \, d\mu \\ &\le \left(\left(\int |f|^p \, d\mu \right)^{1/p} + \int |g|^p \, d\mu \right)^{1/p} \Bigg) \underbrace{\left(\int |f + g|^{(p-1)q} \, d\mu \right)^{1/q}}_{=\left(\int |f+g|^p \, d\mu\right)^{(p-1)/p}} \end{aligned}$$

wobei wir für das zweite Ungleichheitszeichen die Hölder-Ungleichung (Satz 3.20) verwenden. □

[7] Hermann Minkowski, 1864–1909

Beobachtung 3.27. Für $1 \le p \le \infty$ ist $L^p(\mu) = \{[f] : f \in \mathcal{L}^p(\mu)\}$ mit $[f] = \{g : g = f\ \mu\text{-f.ü.}\}$ ein Banachraum, $L^2(\mu)$ ist mit Skalarprodukt $\langle [f],[g]\rangle := \int fg\,d\mu$ ein Hilbertraum.

Bemerkung 3.28. Sei $\mu(S) < \infty$, $1 \le r < s \le \infty$. Es gilt $\mathcal{L}^s(\mu) \subset \mathcal{L}^r(\mu)$ und die Einbettung $\mathcal{L}^s(\mu) \ni f \mapsto f \in \mathcal{L}^r(\mu)$ ist stetig.

Beweis. Fall $s = \infty$: Sei $f \in \mathcal{L}^\infty(\mu)$, so ist

$$\|f\|_r = \Big(\int |f|^r\,d\mu\Big)^{1/r} \le \Big(\big(\|f\|_\infty\big)^r \mu(S)\Big)^{1/r} = \big(\mu(S)\big)^{1/r}\|f\|_\infty.$$

Fall $s < \infty$: Sei $f \in \mathcal{L}^\infty(\mu)$, $p := \frac{s}{r}$ (> 1), $q := \frac{1}{1-1/p}$ $(= \frac{s}{s-r})$, so ist $\frac{1}{p} + \frac{1}{q} = 1$ und

$$\begin{aligned}\|f\|_r &= \Big(\int 1\cdot |f|^r\,d\mu\Big)^{1/r}\\ &\le \left(\Big(\int 1^q\,d\mu\Big)^{1/q}\Big(\int \big(|f|^r\big)^p\,d\mu\Big)^{1/p}\right)^{1/r} = \big(\mu(S)\big)^{1-r/s}\|f\|_s.\end{aligned}$$

mit Hölder-Ungleichung (Satz 3.20). □

3.1 Gleichgradige Integrierbarkeit

Definition 3.29. Sei $1 \le p < \infty$. Ein Teilmenge $\mathcal{F} \subset \mathcal{L}^p(\mu)$ heißt *gleichgradig p-integrierbar* (für $p = 1$ meist einfach *gleichgradig integrierbar*), wenn gilt

$$\inf_{g\in\mathcal{L}^p(\mu),\, g\ge 0}\ \sup_{f\in\mathcal{F}} \int_{\{|f|>g\}} |f|^p\,d\mu = 0$$

Bemerkung 3.30. $\mathcal{F} \subset \mathcal{L}^p(\mu)$ mit $|\mathcal{F}| < \infty$ ist gleichgradig p-integrierbar.

Satz 3.31. *Sei $1 \le p < \infty$, $f_1, f_2, \dots \in \mathcal{L}^p(\mu)$, $f : S \to \overline{\mathbb{R}}$ m.b. Dann sind äquivalent*

i) $f_n \to f$ in $\mathcal{L}^p(\mu)$

ii) $f_n \to f$ im Maß μ und $\{f_n, n \in \mathbb{N}\}$ ist gleichgradig p-integrierbar.

Beweis. „*i)* ⇒ *ii)*“ : Konvergenz in $\mathcal{L}^p(\mu)$ impliziert Konvergenz im Maß (Bem. 3.23, *ii)*), zu zeigen bleibt: $\{f_n : n \in \mathbb{N}\}$ ist gleichgradig p-integrierbar.

Sei $h := 2|f|$ $(\in \mathcal{L}^p(\mu))$, für jede Teilfolge $n_1 < n_2 < \dots \nearrow \infty$ gibt es eine Teilteilfolge $(n'_k)_k$ mit

$$|f_{n'_k}|^p \mathbf{1}_{\{|f_{n'_k}|\le h\}} \underset{k\to\infty}{\longrightarrow} |f|^p \quad \mu\text{-f.ü.}$$

(verwende Lemma 3.24), somit

$$\int_{\{|f_{n'_k}|\le h\}} |f_{n'_k}|^p\,d\mu \underset{k\to\infty}{\longrightarrow} \int |f|^p\,d\mu$$

(gemäß Satz von der dominierten Konvergenz, Satz 3.13) und

$$\int_{\{|f_{n'_k}|>h\}} |f_{n'_k}|^p \, d\mu \le 2^p \underbrace{\int |f_{n'_k} - f|^p \, d\mu}_{\to 0 \text{ für } k\to\infty \text{ n. Vor.}} + 2^p \underbrace{\int |f|^p \mathbf{1}_{\{|f_{n'_k}|>h\}} \, d\mu}_{\to 0 \text{ (dom. Konv.)}}$$

Insbesondere gilt

$$\int_{\{|f_n|>h\}} |f_n|^p \, d\mu \xrightarrow[n\to\infty]{} 0.$$

Zu $\varepsilon > 0$ wähle n_0 mit $\int_{\{|f_n|>h\}} |f_n|^p \, d\mu \le \varepsilon$ für $n > n_0$, setze $g := |f_1| + \cdots + |f_{n_0}| + h$, so gilt

$$\sup_n \int_{\{|f_n|>h\}} |f_n|^p \, d\mu \le \varepsilon$$

„*ii) ⇒ i)*“ : o.E. sei $f_n \to f$ μ-f.ü. (sonst gehe zu μ-f.ü. konvergenter Teilfolge über gemäß Lemma 3.24), es ist

$$\int |f|^p \, d\mu \le \liminf_{n\to\infty} \int |f_n|^p \, d\mu \ \ (< \infty)$$

also $f \in \mathcal{L}^p(\mu)$.

Zeige: $\int |f_n - f|^p \, d\mu \xrightarrow[n\to\infty]{} 0$:

Zu $\varepsilon > 0$, $0 \le g \in \mathcal{L}^p(\mu)$ gem. Voraussetzung setze $h := g + 2|f|$. Wie vorher ist

$$\int_{\{|f_n|\le h\}} |f_n|^p \, d\mu \xrightarrow[n\to\infty]{} \int |f|^p \, d\mu,$$

somit

$$\limsup_{n\to\infty} \left| \int |f_n|^p \, d\mu - \int |f|^p \, d\mu \right| \le \limsup_{n\to\infty} \int_{\{|f_n|>h\}} |f_n|^p \, d\mu \le \varepsilon.$$

Weiter ist

$$0 \le 2^p\big(|f|^p + |f_n|^p\big) - |f - f_n|^p \to 2^{p+1}|f|^p \quad \mu\text{-f.ü.},$$

mit Lemma von Fatou (Satz 3.12) daher

$$2^{p+1} \int |f|^p \, d\mu \le \liminf_{n\to\infty} \left(2^p \int |f|^p \, d\mu + 2^p \int |f_n|^p \, d\mu - \int |f_n - f|^p \, d\mu \right).$$

Somit gilt $\limsup_{n\to\infty} \int |f_n - f|^p \, d\mu = 0$. □

Bemerkung 3.32. Sei $\mu(S) < \infty$, dann ist die Bedingung

$$\inf_{M\in\mathbb{R}_+} \sup_{f\in\mathcal{F}} \int_{\{|f|>M\}} |f| \, d\mu = 0 \tag{3.3}$$

äquivalent zur gleichgradigen Integrierbarkeit (von $\mathcal{F}$ bzgl. μ).

Beweis. Da jedes $M \in \mathbb{R}_+$, aufgefasst als konstante Funktion, in $\mathcal{L}^1(\mu)$ liegt, impliziert (3.3) gleichgradige Integrierbarkeit.

Zeige: Sei $\mathcal{F} \subset \mathcal{L}^1(\mu)$ gleichgradig integrierbar $\Rightarrow$ (3.3) :
Zu $\varepsilon > 0$ wähle $g \in \mathcal{L}^1(\mu)$, $g \geq 0$ mit

$$\int_{\{|f|>g\}} |f| \, d\mu < \frac{\varepsilon}{2} \quad \text{für } f \in \mathcal{F},$$

wähle $M > 0$ mit $\int_{\{|g| \geq M\}} |g| \, d\mu < \frac{\varepsilon}{2}$, somit

$$\int_{\{|f|>M\}} |f| \, d\mu \leq \int_{\{|f|>g\}} |f| \, d\mu + \int \mathbf{1}_{\{g \geq M\}} |g| \, d\mu = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

□

Satz 3.33. *Sei $\mu(S) < \infty$, $\mathcal{F} \subset \mathcal{L}^1(\mu)$.*
$\mathcal{F}$ ist gleichgradig integrierbar g.d.w.

$$\textit{es gibt } H : [0, \infty) \to [0, \infty) \textit{ mit } \lim_{x \to \infty} \frac{H(x)}{x} = \infty \textit{ und } \sup_{f \in \mathcal{F}} \int H(|f|) \, d\mu < \infty.$$

H kann monoton wachsend und konvex gewählt werden.

Beweis. „$\Leftarrow$“ : $K_M := \inf_{x \geq M} \frac{H(x)}{x} \underset{M \to \infty}{\nearrow} \infty$, also

$$\sup_{f \in \mathcal{F}} \int_{\{|f|>M\}} |f| \, d\mu \leq \frac{1}{K_M} \int_{\{|f|>M\}} H(|f|) \, d\mu$$
$$\frac{1}{K_M} \int H(|f|) \, d\mu \underset{M \to \infty}{\longrightarrow} 0$$

„$\Rightarrow$“ : Sei $\mathcal{F}$ gleichgradig integrierbar, wähle (gem. Bem. 3.32) $M_n \nearrow \infty$ mit

$$\forall f \in \mathcal{F} : \quad \int \left(|f| - M_n\right)^+ d\mu \leq \int_{\{|f|>M_n\}} |f| \, d\mu \leq 2^{-n},$$

setze $H(x) := \sum_{n=1}^{\infty} (x - M_n)^+$, für $x \geq 2M_n$ ist

$$\frac{H(x)}{x} \geq \sum_{j=1}^{n} \left(1 - \frac{M_j}{x}\right) \geq \frac{n}{2}, \quad \text{also } \frac{H(x)}{x} \underset{x \to \infty}{\longrightarrow} \infty$$

und

$$\forall f \in \mathcal{F} : \quad \int H(|f|) \, d\mu = \sum_{n=1}^{\infty} \int \left(|f| - M_n\right)^+ d\mu \leq \sum_{n=1}^{\infty} 2^{-n} < \infty.$$

□

Korollar 3.34. *$\mu(S) < \infty$, $p > 1$, $\mathcal{F} \subset \mathcal{L}^p(\mu)$ ($\subset \mathcal{L}^1(\mu)$) mit $\sup_{f \in \mathcal{F}} \|f\|_p < \infty$ („$\mathcal{F}$ ist $\mathcal{L}^p(\mu)$-beschränkt“), so ist $\mathcal{F}$ gleichgradig integrierbar.*

(Wähle $H(x) = x^p$ in Satz 3.33.)

3.2 Parameterintegrale

$(S, \mathscr{A}, \mu)$ Maßraum, U metrischer Raum, $f : S \times U \to \mathbb{R}$

Satz 3.35. *Sei $u_0 \in U$, es gelte*

i) $u \mapsto f(x,u)$ *ist stetig in* u_0 *für* μ*-f.a.* $x \in S$

ii) $x \mapsto f(x,u)$ *ist messbar für alle* $u \in U$

iii) $h : x \mapsto \sup_{u \in U} |f(x,u)|$ *liegt in* $\mathcal{L}^1(\mu)$

Dann ist $F : U \to \mathbb{R}$, $F(u) = \int f(x,u)\,\mu(dx)$ *stetig in* u_0

Beweis. Sei $U \ni u_n \to u_0$:

$$\left|F(u_n) - F(u_0)\right| \le \int \underbrace{|f(x,u_n) - f(x,u_0)|}_{\le 2h(x)}\,\mu(dx) \xrightarrow[n\to\infty]{} 0$$

mit dominierter Konvergenz (Satz 3.13). □

Satz 3.36. $U \subset \mathbb{R}$ *offenes Intervall,* $f : S \times U \to \mathbb{R}$ *mit*

i) $\forall\, u \in U :\ x \mapsto f(x,u)$ *liegt in* $\mathcal{L}^1(\mu)$

ii) Für μ*-f.a.* $x \in S$ *ist* $u \mapsto f(x,u)$ *nach* u *differenzierbar (mit Ableitung* $f'(x,u)$*)*

iii) $h : x \mapsto \sup_{u \in U} |f'(x,u)|$ *liegt in* $\mathcal{L}^1(\mu)$

Dann ist für $u \in U$ $f'(\cdot,u) \in \mathcal{L}^1(\mu)$ *und* $F(u) = \int f(x,u)\,\mu(dx)$ *ist auf* U *diff'bar mit* $F'(u) = \int f'(x,u)\,\mu(dx)$.

Beweis. Sei $U \ni u_n \neq u_0$, $u_n \to u_0$ für $n \to \infty$,

$$g_n(x) := \frac{f(x,u_n) - f(x,u_0)}{u_n - u_0}, \quad \text{also } g_n \to f'(\cdot,u_0)\ \mu\text{-f.ü.}$$

Zwischenwertsatz der Differentialrechnung liefert:

$$\exists\, v_n(x) \in U \text{ mit } g_n(x) = f'\big(x, v_n(x)\big) \quad \text{für } \mu\text{-f.a. } x,$$

somit $|g_n| \le h$ μ-f.ü.

Dominierte Konvergenz (Satz 3.13) liefert $f'(\cdot,u_0) \in \mathcal{L}^1(\mu)$ und

$$\lim_{n\to\infty} \frac{F(u_n) - F(u_0)}{u_n - u_0} = \lim_{n\to\infty} g_n(x)\,\mu(dx) = \int f'(x,u_0)\,\mu(dx)$$

□

Beispiel 3.37 (Laplace-Transformierte)**.** $X \ge 0$ ZV auf W'raum $(\Omega, \mathscr{F}, \mathbb{P})$, $U = [0,\infty)$, $f(x,u) = e^{-ux}$. Dann ist $F(u) := \mathbb{E}\big[e^{-uX}\big]$ ∞-oft diff'bar in $(0,\infty)$ und

$$F^{(k)}(u) = (-1)^k \mathbb{E}\big[X^k e^{-uX}\big]$$

3.3 Zur Dichtetransformation

Satz 3.38 (Dichtetransformationsformel)**.** *$U, V \subset \mathbb{R}^d$ offen, $\psi : U \to V$ ein C^1-Diffeomorphismus, μ Maß auf U mit Dichte f bzgl. dem Lebesgue-Maß λ auf $\mathbb{R}^d$, dann hat das Bildmaß $\nu := \mu \circ \psi^{-1}$ (auf V) die Dichte*

$$g(v) = \frac{1}{\big|\det D\psi\big(\psi^{-1}(v)\big)\big|} f\big(\psi^{-1}(v)\big), \quad v \in V$$

bezüglich λ, wo $D\psi(u) = \Big(\frac{\partial \psi_i}{\partial u_j}(u)\Big)_{i,j=1}^d$ die Jacobi-Matrix (von ψ and der Stelle $u \in U$) ist.

Satz 3.39 (Transformationsformel von Jacobi, „d-dimensionale Substitution“)**.** *$U, V \subset \mathbb{R}^d$ offen, $\varphi : U \to V$ C^1-Diffeomorphismus, dann gilt*

$$\lambda\big(\varphi(B)\big) = \int_B |\det D\varphi| \, d\lambda \quad \textit{für } B \subset U \;\; \textit{Borel-messbar} \tag{3.4}$$

und

$$\int_V f(v)\,\lambda(dv) = \int_U f\big(\varphi(u)\big)\,|\det D\varphi(u)|\,\lambda(du) \quad \textit{für } f \in \mathcal{L}^1(\lambda|_V) \tag{3.5}$$

Satz 3.38 folgt aus Satz 3.39 mit der Wahl $\varphi = \psi^{-1}$: Für $A \in \mathcal{B}(V)$ ist

$$\begin{aligned}
\big(\mu \circ \psi^{-1}\big)(A) = \mu\big(\psi^{-1}(A)\big) &= \int_U \big(\mathbf{1}_{\psi^{-1}(A)} \cdot f\big)(u)\,\lambda(du) \\
&= \int_V \big(\mathbf{1}_{\psi^{-1}(A)} \cdot f\big)(\psi^{-1}(v))\,\big|\det(D\psi^{-1})(v)\big|\,\lambda(dv) \\
&= \int_V \mathbf{1}_A(v) \frac{f\big(\psi^{-1}(v)\big)}{\big|\det D\psi\big(\psi^{-1}(v)\big)\big|}\,\lambda(dv)
\end{aligned}$$

(denn $(D\psi^{-1})(v) = (D\psi)^{-1}(\psi^{-1}(v))$ gemäß dem Satz über die Ableitung der Umkehrfunktion).

Beweise von Satz 3.39 finden sich in Analysis-Lehrbüchern, z.B. G. Kersting und M. Brokate, *Maß und Integral*, S. 107, H. Heuser, *Analysis, Teil 2*, Satz 205.2 ("Substitutions-Regel"), O. Forster, *Analysis 3*, Kap. 9, Satz 1 ("Transformationsformel").

Wir betrachten hier nur eine *Beweisheuristik*.

Zu (3.4): Es genügt, (3.4) für halboffene Quader $B = (a, b]$ mit $a, b \in \mathbb{Q}^d$ und $[a, b] \subseteq U$ zu zeigen (diese bilden einen $\cap$-stabilen Erzeuger von $\mathcal{B}(U)$).

Sei dazu $Q_c := (-c, c]^d$ mit $c > 0$. Da eine differenzierbare Funktion „lokal fast wie eine lineare Funktion aussieht“, gilt

$$\varphi(x + Q_c) \approx \varphi(x) + (D\varphi(x))\,(Q_c) \quad \text{für } x \in U \text{ und kleine Werte von } c.$$

Präziser: Sind $K \subseteq U$ kompakt und $0 < \varepsilon < 1$, so gilt für genügend kleine $c > 0$

$$\forall x \in K: \quad \varphi(x) + (1-\varepsilon)\cdot(D\varphi(x))\,(Q_c) \subseteq \varphi(x + Q_c) \subseteq \varphi(x) + (1+\varepsilon)\cdot(D\varphi(x))\,(Q_c), \tag{3.6}$$

d.h. das Bild des Würfels $x + Q_c$ unter φ lässt sich von von außen und innen unter Benutzung des Parallelotops $(D\varphi(x))(Q_c)$ „einschachteln".

Dann betrachtet man eine Folge von immer feiner werdenden Zerlegungen von $B = (a,b]$ in disjunkte „Kopien" von Q_c und verwendet den Satz von der dominierten Konvergen, um (3.4) für B aus (3.6) zu folgern.

Zu (3.5): Für $f = \mathbf{1}_A$ mit $A = \varphi^{-1}(B)$ folgt dies direkt aus (3.4). Allgemeines $f \geq 0$ kann man in monoton wachsender Weise mittels Linearkombinationen von Indikatorfunktionen approximieren (vgl. Beob. 3.15, „Monotonieprinzip"), schließlich zerlege $\mathcal{L}^1(\lambda|_V) \ni f = f^+ - f^-$.

3.4 Zum Zusammenhang mit dem Riemann-Integral

Erinnerung 3.40 (Riemann-Integral)**.** $I = [a,b]$, $f : I \to \mathbb{R}$ beschränkt, $\mathbf{t} = \{t_0, t_1, \ldots, t_n\}$ mit $a = t_0 < t_1 < \cdots < t_n = b$ Partition von $[a,b]$ mit Feinheit $w(\mathbf{t}) = \max_{1\leq i\leq n}(t_i - t_{i-1})$, $o_j = o_j(f,\mathbf{t}) \coloneqq \sup_{t_{j-1}\leq t\leq t_j} f(t)$, $u_j = u_j(f,\mathbf{t}) \coloneqq \inf_{t_{j-1}\leq t\leq t_j} f(t)$.

$$U_{\mathbf{t}}(f) \coloneqq \sum_{j=1}^{n} u_j \cdot (t_j - t_{j-1}), \quad O_{\mathbf{t}}(f) \coloneqq \sum_{j=1}^{n} o_j \cdot (t_j - t_{j-1})$$

heißen die Unter- und die Obersumme (von f bezgl. der Partition $\mathbf{t}$).

f heißt (eigentlich) *Riemann-integrierbar* (auf I), wenn gilt

$$\lim_{m\to\infty} U_{\mathbf{t}^{(m)}}(f) = \lim_{m\to\infty} O_{\mathbf{t}^{(m)}}(f) =: \int_a^b f(x)\,dx$$

für jede Folge von Partitionen $(\mathbf{t}^{(m)})_{m\in\mathbb{N}}$ mit $w(\mathbf{t}^{(m)}) \to 0$. (Man kann dabei o.E. annehmen, dass $\mathbf{t}^{(m)} \subset \mathbf{t}^{(m+1)}$ gilt, d.h. die $\mathbf{t}^{(m)}$ bilden eine Folge von Verfeinerungen, sonst gehe sukzessive zur gemeinsamen Verfeinerung über.)

Satz 3.41. *$f : [a,b] \to \mathbb{R}$ sei beschränkt, $C \coloneqq \{x \in [a,b] : f$ stetig in $x\}$, $D \coloneqq [a,b] \smallsetminus C$. Dann gilt*

i) $C, D \in \mathcal{B}([a,b])$ und $f \cdot \mathbf{1}_C$ ist Borel-messbar

ii) f ist Riemann-integrierbar (über $[a,b]$) g.d.w. $\lambda(D) = 0$,

in diesem Fall gilt für sein Riemann-Integral

$$\int_a^b f(x)\,dx = \int f \cdot \mathbf{1}_C\,d\lambda \quad \Big(= \int_{[a,b]} f\,d\lambda, \quad \textit{sofern } f \textit{ Borel-messbar}\Big)$$

Beweisskizze. i) Sei $\mathbf{t}^{(m)} = \{t_0^{(m)}, t_1^{(m)}, \ldots, t_{n_m}^{(m)}\}$, $m \in \mathbb{N}$ eine Folge von Verfeinerungen mit $w(\mathbf{t}^{(m)}) \to 0$,

$$g_m \coloneqq \sum_{j=1}^{n_m} u_j(f,\mathbf{t}^{(m)})\mathbf{1}_{(t_{j-1}^{(m)}, t_j^{(m)}]}, \quad h_m \coloneqq \sum_{j=1}^{n_m} o_j(f,\mathbf{t}^{(m)})\mathbf{1}_{(t_{j-1}^{(m)}, t_j^{(m)}]},$$

dann ist $g_m \leq f \leq h_m$ und $U_{\mathbf{t}^{(m)}}(f) = \int g_m \, d\lambda, O_{\mathbf{t}^{(m)}}(f) = \int h_m \, d\lambda$. Die Folge $(g_m)_m$ steigt auf gegen eine Borel-messbare Funktion g, $(h_m)_m$ steigt ab gegen eine Borel-messbare Funktion h, somit gilt $g \leq f \leq h$ auf I, weiter ist

$$\{x \in I : g(x) < h(x)\} \subset D \subset \{x \in I : g(x) < h(x)\} \cup \bigcup_{m=1}^{\infty} \{t_0^{(m)}, t_1^{(m)}, \ldots, t_{n_m}^{(m+1)}\}$$

und man gewinnt daraus die geforderten Messbarkeitseigenschaften.

ii) Es ist

$$\int g \, d\lambda = \lim_{m\to\infty} \int g_m \, d\lambda = \lim_{m\to\infty} U_{\mathbf{t}^{(m)}}(f),$$
$$\int h \, d\lambda = \lim_{m\to\infty} \int h_m \, d\lambda = \lim_{m\to\infty} O_{\mathbf{t}^{(m)}}(f)$$

(mit dominierter Konvergenz), wegen $g \leq h$ also

$$g = h \ \lambda\text{-f.ü.} \iff \lambda(D) = 0 \iff f \text{ Riemann-integrierbar.}$$

Das Beispiel $f = \mathbf{1}_{\mathbb{Q}\cap[0,1]}$ zeigt, dass eine Lebesgue-integrierbare Funktion nicht Riemann-integrierbar sein muss.

Für das uneigentliche Riemann-Integral über die Achse oder eine Halbachse, das man im Sinne eines Grenzwerts bezüglich der Integrationsgrenze definiert (sofern dieser Grenzwert existiert), gilt die analoge Aussage nicht immer: Beispielsweise existiert

$$\int_0^\infty \frac{\sin(x)}{x} \, dx = \lim_{R\to\infty} \int_0^R \frac{\sin(x)}{x} \, dx \ (= \pi/2)$$

im Sinne der uneigentlichen Riemann-Integrale, aber die Funktion $\sin(x)/x$ liegt nicht in $\mathcal{L}^1(\lambda|_{[0,\infty)})$ (denn $\int_{[0,\infty)} |\sin(x)/x| \, \lambda(dx) = \infty$).

Kapitel 4

Gesetz der großen Zahlen

Erinnerung 4.1 (Varianz und Kovarianz)**.** $(\Omega, \mathscr{F}, \mathbb{P})$ W'raum, $X, Y \in \mathcal{L}^2(\mathbb{P})$.

$$\mathrm{Var}[X] := \mathbb{E}\big[(X - \mathbb{E}[X])^2\big] = \mathbb{E}\big[X^2\big] - \big(\mathbb{E}[X]\big)^2 \;\; (\geq 0), \quad \text{die } \textit{Varianz} \text{ von } X$$

($\sqrt{\mathrm{Var}[X]}$ ist die *Streuung* oder *Standardabweichung* von X)

$$\begin{aligned}\mathrm{Cov}[X,Y] &:= \mathbb{E}\big[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\big] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \;\; \big(= \mathrm{Cov}[X,Y]\big), \qquad \text{die } \textit{Kovarianz} \text{ von } X \text{ und } Y\end{aligned}$$

Für $X, Y, Z \in \mathcal{L}^2(\mathbb{P})$, $a, b \in \mathbb{R}$ gilt

$$\mathrm{Cov}[aX + bY, Z] = a\mathrm{Cov}[X,Z] + b\mathrm{Cov}[Y,Z]$$

(die Kovarianz ist eine symmetrische Bilinearform), insbesondere gilt

$$\mathrm{Var}[aX] = \mathrm{Cov}[aX, aX] = a^2\mathrm{Var}[X].$$

X und Y heißen *unkorreliert*, wenn $\mathrm{Cov}[X,Y] = 0$; für paarweise unkorrelierte $X_1, \dots, X_n \in \mathcal{L}^2(\mathbb{P})$ gilt

$$\mathrm{Var}[X_1 + \dots + X_n] = \sum_{i,j=1}^{n} \mathrm{Cov}[X_i, X_j] = \sum_{i=1}^{n} \mathrm{Var}[X_i].$$

Erinnerung 4.2. $X, Y \in \mathcal{L}^2(\mathbb{P})$ unabhängig $\implies$ X, Y unkorreliert.
Die Umkehrung gilt i.A. nicht.

Beweis. Falls X und Y jeweils höchstens abzählbar viele Werte annehmen:

$$\begin{aligned}\mathbb{E}[XY] &= \sum_{x,y} xy\mathbb{P}(X = x, Y = y) = \sum_{x,y} xy\mathbb{P}(X = x)\mathbb{P}(Y = y) \\ &= \Big(\sum_{x} x\mathbb{P}(X = x)\Big)\Big(\sum_{y} y\mathbb{P}(Y = y)\Big) = \mathbb{E}[X]\,\mathbb{E}[Y],\end{aligned}$$

also $\mathrm{Cov}[X,Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$.

Allgemeiner Fall: o.E. seien $X, Y \geq 0$ (sonst zerlege $X = X^+ - X^-$, $Y = Y^+ - Y^-$), es ist

$$\mathbb{E}[XY] = \lim_{n\to\infty} \mathbb{E}\big[2^{-n}\lfloor 2^n X\rfloor\, 2^{-n}\lfloor 2^n Y\rfloor\big] = \lim_{n\to\infty} \mathbb{E}\big[2^{-n}\lfloor 2^n X\rfloor\big]\mathbb{E}\big[2^{-n}\lfloor 2^n Y\rfloor\big] = \mathbb{E}[X]\mathbb{E}[Y]$$

(verwende monotone Konvergenz, Satz 3.11).

Sei X symmetrisch verteilt mit $0 < \mathbb{E}[X^4] < \infty$, dann sind X und $Y := X^2$ wegen $\mathbb{E}[XY] = \mathbb{E}[X^3] = 0 = \mathbb{E}[X] = \mathbb{E}[X]\mathbb{E}[Y]$ unkorreliert, aber nicht unabhängig.

□

Erinnerung 4.3 (Chebyshev[1]-Ungleichung)**.** Für $X \in \mathcal{L}^2(\mathbb{P})$ und $a > 0$ gilt

$$\mathbb{P}\big(|X - \mathbb{E}[X]| \geq a\big) \leq \frac{\operatorname{Var}[X]}{a^2}$$

(denn $a^2\mathbb{P}\big(|X - \mathbb{E}[X]| \geq a\big) = a^2\mathbb{P}\big(|X - \mathbb{E}[X]|^2 \geq a^2\big) \leq \mathbb{E}\big[|X - \mathbb{E}[X]|^2\big]$ gemäß Markov-Ungleichung, Beob 3.3).

Erinnerung 4.4 (Schwaches Gesetz der großen Zahlen via Chebyshev-Ungleichung)**.** $X_1, X_2, \dots \in \mathcal{L}^2(\mathbb{P})$ paarweise unkorrelierte ZVn mit $\mathbb{E}[X_i] = 0$, $\sup_{i\in\mathbb{N}} \mathbb{E}[X_i^2] =: c < \infty$, dann gilt für $S_n := X_1 + \dots + X_n$

$$\frac{1}{n} S_n \underset{n\to\infty}{\longrightarrow} 0 \quad \mathbb{P}\text{-stochastisch},$$

denn für $\varepsilon > 0$ gilt

$$\mathbb{P}\big(|S_n| > \varepsilon n\big) \leq \frac{1}{\varepsilon^2 n^2}\operatorname{Var}[S_n] = \frac{1}{\varepsilon^2 n^2}\sum_{i=1}^n \operatorname{Var}[X_i] \leq \frac{c}{\varepsilon^2 n} \underset{n\to\infty}{\longrightarrow} 0$$

Erinnerung 4.5. Seien $X_1, X_2, \dots \in \mathcal{L}^2(\mathbb{P})$ paarweise unabhängige ZVn mit $\sup_{i\in\mathbb{N}} \mathbb{E}[X_i^2] =: c < \infty$, es gelte $\mathbb{E}[X_n^+] \to \mu_+ \in \mathbb{R}$, $\mathbb{E}[X_n^-] \to \mu_- \in \mathbb{R}$. Dann gilt für $S_n := X_1 + \dots + X_n$

$$\frac{1}{n} S_n \underset{n\to\infty}{\longrightarrow} \mu := \mu_+ - \mu_- \quad \text{fast sicher.}$$

Beweis. O.E. seien $X_i \geq 0$ (sonst betrachte X_i^+ und X_i^- separat); setze $\mu_i := \mathbb{E}[X_i]$, n. Vor. gilt $\mu_i \to \mu \in \mathbb{R}$ für $i \to \infty$.

Sei $a > 1$, $k_n := \lceil a^n \rceil$

$$\begin{aligned}\sum_{n=1}^\infty \mathbb{P}\Bigg(\Big|\frac{1}{k_n}\sum_{i=1}^{k_n}(X_i - \mu_i)\Big| > a^{-n/4}\Bigg) &\leq \sum_{n=1}^\infty a^{n/2}\operatorname{Var}\Big[\frac{1}{k_n}\sum_{i=1}^{k_n}(X_i - \mu_i)\Big] = \sum_{n=1}^\infty a^{n/2}\frac{1}{k_n^2}\sum_{i=1}^{k_n}\operatorname{Var}[X_i] \\ &\leq \sum_{n=1}^\infty a^{n/2}\frac{c}{k_n} \leq c\sum_{n=1}^\infty a^{-n/2} < \infty,\end{aligned}$$

also gibt es mit Borel-Cantelli-Lemma (Lemma 2.5) ein zufälliges N_0 mit

$$\Big|\frac{1}{k_n}\sum_{i=1}^{k_n}(X_i - \mu_i)\Big| = \Big|\frac{1}{k_n}S_{k_n} - \underbrace{\frac{1}{k_n}\sum_{i=1}^{k_n}\mu_i}_{\to\mu}\Big| \leq a^{-n/4} \quad \text{für } n \geq N_0$$

[1] Pafnuty Lvovich Chebyshev, 1821–1894.

und insbesondere gilt $\frac{1}{k_n} S_{k_n} \xrightarrow[n\to\infty]{} \mu$ fast sicher.

Sei $b > a$ (> 1), dann ist $k_{n+1} \le b\, k_n$ (für n gen. groß), also für $k_n \le m \le k_{n+1}$

$$\frac{1}{b} \underbrace{\frac{S_{k_n}}{k_n}}_{\to\mu} \le \frac{S_{k_n}}{k_{n+1}} \le \frac{S_m}{m} \le \frac{S_{k_{n+1}}}{k_n} \le b \underbrace{\frac{S_{k_{n+1}}}{k_{n+1}}}_{\to\mu}$$

somit

$$\Big(\frac{1}{b} - 1\Big)\mu \le \liminf_{m\to\infty} \Big(\frac{S_m}{m} - \mu\Big) \le \limsup_{m\to\infty} \Big(\frac{S_m}{m} - \mu\Big) \le (b-1)\mu$$

mit $b \downarrow 1$ folgt die Beh. $\square$

Satz 4.6 (Starkes Gesetz der großen Zahlen)**.** $X_1, X_2, \dots \in \mathcal{L}^1(\mathbb{P})$ *seien identisch verteilt und paarweise unabhängig, dann gilt*

$$\frac{1}{n}\big(X_1 + \dots + X_n\big) \xrightarrow[n\to\infty]{} \mu := \mathbb{E}[X_1] \quad \textit{fast sicher.} \tag{4.1}$$

Lemma 4.7. $X_1, X_2, \dots$ *wie in Satz 4.6,* $Y_n := X_n \mathbf{1}_{\{|X_n| \le n\}}$, $T_n := Y_1 + \dots + Y_n$. *Dann gilt*

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}[Y_n^2]}{n^2} \le \mathbb{E}\big[|X_1|\big] + 1$$

und $\frac{1}{n} T_n \xrightarrow[n\to\infty]{} \mu$ *f.s. impliziert* (4.1).

Beweis. Für $x > 0$ ist

$$\sum_{n=1}^{\infty} \frac{x^2}{n^2} \mathbf{1}_{\{x \le n\}} = \frac{x^2}{(\lceil x\rceil)^2} + \sum_{n=\lceil x\rceil+1}^{\infty} x^2 \frac{1}{n^2} \le 1 + x^2 \int_{\lceil x\rceil}^{\infty} \frac{1}{y^2}\, dy = 1 + \frac{x^2}{\lceil x\rceil} \le 1 + x,$$

somit

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{\mathbb{E}[Y_n^2]}{n^2} &= \sum_{n=1}^{\infty} \mathbb{E}\Big[\frac{X_n^2}{n^2} \mathbf{1}_{\{|X_n| \le n\}}\Big] \\ &= \sum_{n=1}^{\infty} \mathbb{E}\Big[\frac{X_1^2}{n^2} \mathbf{1}_{\{|X_1| \le n\}}\Big] = \mathbb{E}\Big[\sum_{n=1}^{\infty} \frac{X_1^2}{n^2} \mathbf{1}_{\{|X_1| \le n\}}\Big] \le \mathbb{E}\big[1 + |X_1|\big]. \end{aligned}$$

Es ist

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n \ne Y_n) = \sum_{n=1}^{\infty} \mathbb{P}\big(|X_n| > n\big) = \sum_{n=1}^{\infty} \mathbb{P}\big(|X_1| > n\big) \le \mathbb{E}\big[|X_1|\big] < \infty$$

mit Borel-Cantelli-Lemma (Lemma 2.5) gibt es ein zufälliges N_0, so dass $X_n = Y_n$ für $n \ge N_0$, insbesondere

$$\frac{1}{n} \sum_{i=1}^{n} (X_i - Y_i) \xrightarrow[n\to\infty]{} 0 \quad \text{f.s.}$$

Wegen $\frac{1}{n}\big(X_1 + \dots + X_n\big) = \frac{1}{n} T_n + \frac{1}{n} \sum_{i=1}^{n} (X_i - Y_i)$ impliziert dies die zweite Behauptung. $\square$

Beweis von Satz 4.6. O.E. seien $X_n \geq 0$ (sonst zerlege in $X_n = X_n^+ - X_n^-$).

Sei $a > 1$, $k_n := \lceil a^n \rceil$ $(\geq a^n/2)$. Für $m \in \mathbb{N}$ ist

$$\sum_{n:k_n \geq m} k_n^{-2} \leq \sum_{n=\lfloor(\log m)/(\log a)\rfloor}^{\infty} \frac{4}{a^{2n}} = \frac{4}{1-a^{-2}} a^{-2\lfloor(\log m)/(\log a)\rfloor} \leq \frac{4a^2}{1-a^{-2}} m^{-2}$$

Sei T_n wie in Lemma 4.7. Für $\delta > 0$ ist

$$\begin{aligned}\sum_{n=1}^{\infty} \mathbb{P}\Big(\big|T_{k_n} - \mathbb{E}[T_{k_n}]\big| > \delta k_n\Big) &\leq \delta^{-2} \sum_{n=1}^{\infty} k_n^{-2} \sum_{m=1}^{k_n} \operatorname{Var}[Y_m] \\ &= \delta^{-2} \sum_{m=1}^{\infty} \operatorname{Var}[Y_m] \sum_{n:k_n \geq m} k_n^{-2} \leq \delta^{-2} \frac{4a^2}{1-a^{-2}} \sum_{m=1}^{\infty} \frac{\mathbb{E}[Y_m^2]}{m^2} < \infty\end{aligned}$$

mit Borel-Cantelli also

$$\limsup_{n\to\infty} \left| \frac{T_{k_n} - \mathbb{E}[T_{k_n}]}{k_n} \right| \leq \delta \quad \text{f.s.}$$

(für jedes $\delta > 0$), zudem gilt $\mathbb{E}[Y_n] = \mathbb{E}[X_n \mathbf{1}_{\{|X_n| \leq n\}}] = \mathbb{E}[X_1 \mathbf{1}_{\{|X_1| \leq n\}}] \to \mathbb{E}[X_1] = \mu$ (dominierte Konvergenz), also auch $\mathbb{E}[T_{k_n}]/k_n \to \mu$, d.h.

$$\frac{T_{k_n}}{k_n} \underset{n\to\infty}{\longrightarrow} \mu \quad \text{f.s.}$$

Wie im Beweis von Erinnerung 4.5 folgt $T_m/m \to \mu$ f.s. für $m \to \infty$, mit Lemma 4.7 folgt die Beh. $\square$

Kapitel 5

Produktmaße und Übergangskerne

Lemma 5.1. *Seien* $(S_1, \mathscr{A}_1)$ *und* $(S_2, \mathscr{A}_2)$ *messbare Räume,* $f : S_1 \times S_2 \to \overline{\mathbb{R}}$ $(\mathscr{A}_1 \otimes \mathscr{A}_2)$-$(\mathcal{B}(\mathbb{R})$-$)$*messbar,* $A \in \mathscr{A}_1 \otimes \mathscr{A}_2$. *Für* $\widetilde{x}_1 \in S_1, \widetilde{x}_2 \in S_2$ *gilt*

$$\begin{aligned}
&A_{\widetilde{x}_1} := \left\{x_2 \in S_2 : (\widetilde{x}_1, x_2) \in A\right\} \in \mathscr{A}_2 \\
&A_{\widetilde{x}_2} := \left\{x_1 \in S_1 : (x_1, \widetilde{x}_2) \in A\right\} \in \mathscr{A}_1 \\
&f_{\widetilde{x}_1} : S_2 \to \overline{\mathbb{R}},\ \ x_2 \mapsto f_{\widetilde{x}_1}(x_2) = f(\widetilde{x}_1, x_2) \quad \textit{ist } \mathscr{A}_2\text{-}\mathcal{B}(\mathbb{R})\textit{-messbar} \\
&f_{\widetilde{x}_2} : S_1 \to \overline{\mathbb{R}},\ \ x_1 \mapsto f_{\widetilde{x}_2}(x_1) = f(x_1, \widetilde{x}_2) \quad \textit{ist } \mathscr{A}_1\text{-}\mathcal{B}(\mathbb{R})\textit{-messbar}
\end{aligned}$$

Sei $f \geq 0$, μ_i σ*-endliches Maß auf* $(S_i, \mathscr{A}_i)$ *(für* $i = 1, 2$*).*

$$\begin{aligned}
I_1 : x_2 &\mapsto \int_{S_1} f(x_1, x_2)\, \mu_1(dx_1) \quad \textit{ist } \mathscr{A}_2\text{-}\mathcal{B}(\mathbb{R})\textit{-messbar} \\
I_2 : x_1 &\mapsto \int_{S_2} f(x_1, x_2)\, \mu_2(dx_2) \quad \textit{ist } \mathscr{A}_1\text{-}\mathcal{B}(\mathbb{R})\textit{-messbar}
\end{aligned} \tag{5.1}$$

Beweis. $\pi_i : S_1 \times S_2 \to S_i$ $(i = 1, 2)$ seien die kanonischen Koordinatenprojektionen, $\widetilde{x}_1 \in S_1$, $\iota : S_2 \to S_1 \times S_2$, $S_2 \ni x_2 \mapsto \iota(x_2) = (\widetilde{x}_1, x_2)$ die Einbettung.

$\pi_1 \circ \iota$ ist konstant $(\equiv \widetilde{x}_1)$, insbesondere $\mathscr{A}_1$-m.b.; $\pi_2 \circ \iota = \mathrm{Id}_{S_2}$ ist $\mathscr{A}_2$-m.b., also mit Lemma 1.43: ι ist $\mathscr{A}_2$-$(\mathscr{A}_1 \otimes \mathscr{A}_2)$-messbar. Demnach: $A_{\widetilde{x}_1} = \iota^{-1}(A) \in \mathscr{A}_2$, $f_{\widetilde{x}_1} = f \circ \iota$ ist $\mathscr{A}_2$-messbar.

Der Fall mit fixiertem $\widetilde{x}_2 \in S_2$ geht analog.

Seien zunächst μ_1, μ_2 jeweils endliche Maße. (5.1) gilt für $f = \mathbf{1}_{A_1 \times A_2}$ mit $A_i \in \mathscr{A}_i$, denn dann ist $I_1(x_2) = \mu_1(A_1)\mathbf{1}_{A_2}(x_2)$, $I_2(x_1) = \mu_2(A_2)\mathbf{1}_{A_1}(x_1)$. Daher enthält

$$\mathscr{D} := \left\{A \in \mathscr{A}_1 \otimes \mathscr{A}_2 :\ (5.1) \text{ gilt für } f = \mathbf{1}_A\right\}$$

zumindest alle Zylindermengen $A = A_1 \times A_2$. $\mathscr{D}$ *ist* ein Dynkin-System, da $\{A_1 \times A_2 : A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2\}$ ein $\cap$-stabiler Erzeuger von $\mathscr{A}_1 \otimes \mathscr{A}_2$ ist, folgt mit Satz 1.14, dass $\mathscr{D} = \mathscr{A}_1 \otimes \mathscr{A}_2$.

Folglich:

(5.1) gilt für alle Elementarfunktionen.

Ein allgemeines messbares $f \geq 0$ kann wachsend mit Elementarfunktionen approximiert werden, z.B. via

$$f = \lim_{n\to\infty} 2^{-n}\lfloor 2^n f\rfloor \wedge n,$$

mit montoner Konvergenz (Satz 3.11) und der Tatsache, dass der punktweise Limes messbarer Funktionen messbar ist (Lemma 1.45), folgt: (5.1) gilt allgemein.

Falls μ_i (nur) σ-endlich sind: Seien $B_n \nearrow S_1, C_n \nearrow S_2$ mit $\mu_1(B_n), \mu_2(C_n) < \infty$, so gilt (5.1) für $f_n(x_1,x_2) := \mathbf{1}_{B_n}(x_1)\mathbf{1}_{C_n}(x_2)f(x_1,x_2) \nearrow f(x_1,x_2)$. □

Satz 5.2. *Seien $(S_i, \mathscr{A}_i, \mu_i)$, $i = 1,2$ σ-endliche Maßräume, $S = S_1 \times S_2$, $\mathscr{A} = \mathscr{A}_1 \otimes \mathscr{A}_2$.*

Es gibt genau ein Maß μ (geschrieben $\mu = \mu_1 \otimes \mu_2$) auf $(S, \mathscr{A})$ mit

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2), \quad A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_1. \tag{5.2}$$

μ heißt das Produktmaß *von μ_1 und μ_2; μ ist σ-endlich.*

Für $f \geq 0$ $\mathscr{A}$-$\mathcal{B}(\overline{\mathbb{R}})$-messbar oder $f \in \mathcal{L}^1(\mu_1 \otimes \mu_2)$ gilt

$$\begin{aligned}\int_S f\, d(\mu_1 \otimes \mu_2) &= \int_{S_2} \left(\int_{S_1} f(x_1,x_2)\, \mu_1(dx_1) \right) \mu_2(dx_2) \\ &= \int_{S_1} \left(\int_{S_2} f(x_1,x_2)\, \mu_2(dx_2) \right) \mu_1(dx_1)\end{aligned} \tag{5.3}$$

Beweis. Eindeutigkeit und σ-Endlichkeit von μ ✓

Zur Existenz: Setze

$$\mu(A) := \int_{S_2} \left(\int_{S_1} \mathbf{1}_A(x_1,x_2)\, \mu_1(dx_1) \right) \mu_2(dx_2), \quad A \in \mathscr{A}$$

(ist wohldefiniert: gemäß Lemma 5.1 ist $x_2 \mapsto \int_{S_1} \mathbf{1}_A(x_1,x_2)\, \mu_1(dx_1) \geq 0$ eine messbare Funktion von x_2).

μ ist ein Maß: Offenbar $\mu(\varnothing) = 0$, seien $A_1, A_2, \dots \in \mathscr{A}$ paarweise disjunkt, $A := \dot\bigcup_n A_n$.

$$\begin{aligned}\mu(A) &= \int_{S_2} \left(\int_{S_1} \sum_{n=1}^{\infty} \mathbf{1}_{A_n}(x_1,x_2)\, \mu_1(dx_1) \right) \mu_2(dx_2) \\ &= \sum_{n=1}^{\infty} \int_{S_2} \left(\int_{S_1} \mathbf{1}_{A_n}(x_1,x_2)\, \mu_1(dx_1) \right) \mu_2(dx_2) = \sum_{n=1}^{\infty} \mu(A_n)\end{aligned}$$

und für $A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2$ ist

$$\begin{aligned}\mu(A_1 \times A_2) &= \int_{S_2} \left(\int_{S_1} \mathbf{1}_{A_1}(x_1)\mathbf{1}_{A_2}(x_2)\, \mu_1(dx_1) \right) \mu_2(dx_2) \\ &= \int_{S_2} \mathbf{1}_{A_2}(x_2) \Bigg(\underbrace{\int_{S_1} \mathbf{1}_{A_1}(x_1)\, \mu_1(dx_1)}_{=\mu_1(A_1)} \Bigg) \mu_2(dx_2) = \mu_1(A_1)\mu_2(A_2).\end{aligned}$$

Zu (5.3) :

$$\widetilde{\mu}(A) := \int_{S_1} \left(\int_{S_2} \mathbf{1}_A(x_1, x_2)\, \mu_2(dx_2) \right) \mu_1(dx_1), \quad A \in \mathscr{A}$$

ist ebenfalls ein Maß, das (5.2) erfüllt, also $\mu = \widetilde{\mu}$.

Demnach gilt (5.3) für $f = \mathbf{1}_A$ mit $A \in \mathscr{A}$.

$$\{f \geq 0 : f \text{ m.b., (5.3) gilt für } f\}$$

erfüllt die Voraussetzungen von Beob. 3.15, also gilt (5.3) für jedes messbare $f \geq 0$ (möglicherweise als $+\infty = +\infty = +\infty$).

Sei nun $f = f^+ - f^- \in \mathcal{L}^1(\mu)$.

$$C_1 := \left\{ x_1 \in S_1 : \int_{S_2} f^+(x_1, x_2)\, \mu_2(dx_2) = \infty \text{ oder } \int_{S_2} f^-(x_1, x_2)\, \mu_2(dx_2) = \infty \right\}$$

erfüllt $\mu_1(C_1) = 0$,

$$C_2 := \left\{ x_2 \in S_2 : \int_{S_1} f^+(x_1, x_2)\, \mu_1(dx_1) = \infty \text{ oder } \int_{S_1} f^-(x_1, x_2)\, \mu_1(dx_1) = \infty \right\}$$

erfüllt $\mu_2(C_2) = 0$, somit

$$\mu(M) = 0 \text{ für } M := C_1 \times S_2 \cup S_1 \times C_2 = (C_1^c \times C_2^c)^c$$

und

$$\begin{aligned}
\int_S f\, d(\mu_1 \otimes \mu_2) &= \int_S \mathbf{1}_{M^c}(f^+ - f^-)\, d(\mu_1 \otimes \mu_2) \\
&= \int_{S_1} \mathbf{1}_{C_1^c} \int_{S_2} \mathbf{1}_{C_2^c}(f^+ - f^-)\, d\mu_2\, d\mu_1 = \int_{S_1} \int_{S_2} (f^+ - f^-)\, d\mu_2\, d\mu_1 \\
&= \int_{S_2} \mathbf{1}_{C_2^c} \int_{S_1} \mathbf{1}_{C_1^c}(f^+ - f^-)\, d\mu_1\, d\mu_2 = \int_{S_2} \int_{S_1} (f^+ - f^-)\, d\mu_1\, d\mu_2.
\end{aligned}$$

□

Satz 5.3. *$(S_i, \mathscr{A}_i, \mu_i)$, $i = 1, 2, \dots, n$ σ-endliche Maßräume, $S := \times_{i=1}^n S_i$, $\mathscr{A} := \bigotimes_{i=1}^n \mathscr{A}_i$. Es gibt genau ein Maß μ auf $(S, \mathscr{A})$ mit*

$$\mu(A_1 \times \cdots \times A_n) = \prod_{i=1}^n \mu_i(A_i), \quad \textit{für } A_i \in \mathscr{A}_i, i = 1, \dots, n.$$

μ heißt das Produktmaß *von $\mu_1, \dots, \mu_n$, geschrieben $\mu = \mu_1 \otimes \cdots \otimes \mu_n = \bigotimes_{i=1}^n \mu_i$.*

Falls $(S_i, \mathscr{A}_i, \mu_i) = (S_1, \mathscr{A}_1, \mu_1)$ für $i = 2, \dots, n$, so schreibt man auch $\mu = \mu_1^{\otimes n}$.

Beweis. Induktiv: Für $n = 2$ ist dies Satz 5.2. Sei $n > 2$,

$$\widetilde{S} := S_1 \times \cdots \times S_{n-1}, \quad \widetilde{\mathscr{A}} := \mathscr{A}_1 \otimes \cdots \otimes \mathscr{A}_{n-1}.$$

Das gewünschte Maß $\widetilde{\mu} = \mu_1 \otimes \cdots \otimes \mu_{n-1}$ auf $(\widetilde{S}, \widetilde{\mathscr{A}})$ existiert nach Ind.ann. Dann ist $S = \widetilde{S} \times S_n$ und das gewünschte $\mu = \widetilde{\mu} \otimes \mu_n$ existiert nach Satz 5.2. □

5.1 (Übergangs-)Kerne

Definition 5.4. Seien $(S_1, \mathscr{A}_1)$ und $(S_2, \mathscr{A}_2)$ messbare Räume. $\kappa: S_1 \times \mathscr{A}_2 \to [0, \infty]$ *heißt (σ-)endlicher Kern von* $(S_1, \mathscr{A}_1)$ *nach* $(S_2, \mathscr{A}_2)$, falls gilt

i) Für alle $A_2 \in \mathscr{A}_2$ gilt: $S_1 \ni x \mapsto \kappa(x, A_2)$ ist $(\mathscr{A}_1 - \mathcal{B}(\mathbb{R})$-)messbar.

ii) Für alle $x \in S_1$ gilt: $\kappa(x, \cdot)$ ist ein (σ-)endliches Maß auf $(S_2, \mathscr{A}_2)$.

κ heißt *stochastischer Kern* oder *Markov-Kern*, wenn in *ii)* gefordert wird, dass $\kappa(x, \cdot)$ ein W'maß ist. κ heißt *sub-stochastisch* oder *sub-Markov*, wenn $\kappa(x, S_2) \leq 1$ für alle $x \in S_1$ gilt.

Beispiel.

i) Sei $S_1 = S_2 = S$ höchstens abzählbar, $\mathscr{A}_1 = \mathscr{A}_2 = 2^S$ und $(p_{xy})_{x,y \in S}$ eine stochastische Matrix. Dann ist $\kappa(x, A) := \sum_{y \in A} p_{xy}$ ein stochastischer Kern von S nach S.

ii) Sei $S_1 = S_2 = \mathbb{R}$, $\mathscr{A}_1 = \mathscr{A}_2 = \mathcal{B}(\mathbb{R})$ und ν ein W'maß auf $\mathbb{R}$. Dann ist $\kappa(x, A) := (\delta_x * \nu)(A) = \nu(A - x)$ ein stochastischer Kern.

   (Interpretation: $\kappa(x, \cdot)$ beschreibt einen zufälligen Sprung gemäß ν von x aus.)

Lemma 5.5. *κ endlicher Übergangskern von $(S_1, \mathscr{A}_1)$ nach $(S_2, \mathscr{A}_2)$, $f : S_1 \times S_2 \to [0, \infty]$ $\mathscr{A}_1 \otimes \mathscr{A}_2$-$\mathcal{B}(\mathbb{R})$-messbar. Dann definiert*

$$I_f : S_1 \to [0, \infty], \quad I_f(x_1) = \int_{S_2} f(x_1, x_2)\, \kappa(x_1, dx_2)$$

eine $\mathscr{A}_1$(-$\mathcal{B}(\mathbb{R})$)-messbare Abbildung.

Beweis. Nach Lemma 5.1 ist $x_2 \mapsto f(x_1, x_2)$ $\mathscr{A}_2$-messsbar, d.h. das Integral ist erklärt.

$$\mathscr{D} := \left\{ A \in \mathscr{A}_1 \otimes \mathscr{A}_2 : I_{\mathbf{1}_A} \text{ ist } \mathscr{A}_1\text{-messsbar } \right\}$$

umfasst die Zylindermengen $\mathscr{Z} = \{A_1 \times A_2 : A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2\}$, *ist* ein Dynkin-System , da $\mathscr{Z}$ ein $\cap$-stabiler Erzeuger von $\mathscr{A}_1 \otimes \mathscr{A}_2$ ist, folgt mit Satz 1.14, dass $\mathscr{D} = \mathscr{A}_1 \otimes \mathscr{A}_2$, d.h. die Behauptung gilt für jedes $f = \mathbf{1}_A$, $A \in \mathscr{A}_1 \otimes \mathscr{A}_2$ und daher auch für Elementarfunktionen.

Ein allgemeines messbares $f \geq 0$ kann wachsend mit Elementarfunktionen approximiert werden, z.B. via

$$f = \lim_{n \to \infty} 2^{-n} \lfloor 2^n f \rfloor \wedge n,$$

mit montoner Konvergenz (Satz 3.11) und der Tatsache, dass der punktweise Limes messbarer Funktionen messbar ist (Lemma 1.45), folgt der allgemeine Fall. □

Satz und Definition 5.6 (Produkt von Kernen, „zweistufiges Experiment"). *$(S_i, \mathscr{A}_i)$, $i = 0, 1, 2$ m.b. Räume, κ_1 endlicher Kern von $(S_0, \mathscr{A}_0)$ nach $(S_1, \mathscr{A}_1)$, κ_2 endlicher Kern von $(S_0 \times S_1, \mathscr{A}_0 \otimes \mathscr{A}_1)$ nach $(S_2, \mathscr{A}_2)$. Dann ist $\kappa_1 \otimes \kappa_2 : S_0 \times \mathscr{A}_1 \otimes \mathscr{A}_2 \to [0, \infty]$,*

$$(\kappa_1 \otimes \kappa_2)(x_0, A) = \int_{S_1} \int_{S_2} \mathbf{1}_A(x_1, x_2)\, \kappa_2\big((x_0, x_1), dx_2\big)\, \kappa_1(x_0, dx_1)$$

ein σ-endlicher Kern von $(S_0, \mathscr{A}_0)$ nach $(S_1 \times S_2, \mathscr{A}_1 \otimes \mathscr{A}_2)$, $\kappa_1 \otimes \kappa_2$ heißt das Produkt von κ_1 und κ_2.

Sind κ_1 und κ_2 (sub-)stochastisch, so auch $\kappa_1 \otimes \kappa_2$.

Analog definieren wir $\kappa_1 \otimes \kappa_2$, wenn κ_2 Kern von $(S_1, \mathscr{A}_1)$ nach $(S_2, \mathscr{A}_2)$, indem wir κ_2 formal als Kern von $(S_0 \times S_1, \mathscr{A}_0 \otimes \mathscr{A}_1)$ nach $(S_2, \mathscr{A}_2)$ auffassen, der nicht von der S_0-Koordinate abhängt.

Beweis. Sei $A \in \mathscr{A}_1 \otimes \mathscr{A}_2$:

$$f_A : S_0 \times S_1 \ni (x_0, x_1) \mapsto \int_{S_2} \mathbf{1}_A(x_1, x_2)\, \kappa_2\big((x_0, x_1), dx_2\big)$$

ist $\mathscr{A}_0 \otimes \mathscr{A}_1$-m.b. nach Lemma 5.5, also ist auch

$$S_0 \ni x_0 \mapsto (\kappa_1 \otimes \kappa_2)\big(x_0, A\big) = \int_{S_1} f_A(x_0, x_1)\, \kappa_1(x_0, dx_1)$$

$\mathscr{A}_0$-m.b. (wiederum nach Lemma 5.5).

Für $x_0 \in S_0$ ist

$$\mathscr{A}_1 \otimes \mathscr{A}_2 \ni A \mapsto (\kappa_1 \otimes \kappa_2)\big(x_0, A\big) \ \ (\geq 0)$$

σ-additiv gem. Satz von der monotonen Konvergenz.

Sei $A_{x_0,n} := \big\{x_1 \in S_1 : \kappa_2\big((x_0, x_1), S_2\big) \leq n\big\} \nearrow_{n\to\infty} S_1$ und

$$(\kappa_1 \otimes \kappa_2)\big(x_0, A_{x_0,n} \times S_2\big) \leq n\kappa_1(x_0, A_{x_0,n}),$$

d.h. $\kappa_1 \otimes \kappa_2$ ist σ-endlich. □

Korollar 5.7. $(S_1, \mathscr{A}_1, \mu)$ *Maßraum mit $\mu(S_1) < \infty$, $(S_2, \mathscr{A}_2)$ messbarer Raum, κ endlicher Kern von S_1 nach S_2.*

Dann gibt es ein eindeutiges σ-endliches Maß $\mu \otimes \kappa$ (auf $(S_1 \times S_2, \mathscr{A}_1 \otimes \mathscr{A}_2)$) mit

$$\mu \otimes \kappa(A_1 \times A_2) = \int_{A_1} \kappa(x_1, A_2)\, \mu(dx_1), \quad A_1 \in \mathscr{A}_1,\ A_2 \in \mathscr{A}_2$$

Ist κ stochastisch und μ ein W'maß, so ist auch $\mu \otimes \kappa$ ein W'maß.

Beweis. Setze $S_0 = \{0\}$, $\mathscr{A}_0 = \big\{\varnothing, \{0\}\big\}$, $\kappa_1(0, A_1) = \mu(A_1)$ und $\kappa_2 = \kappa$ in Satz 5.6. □

Korollar und Definition 5.8. $n \in \mathbb{N}$, $(S_i, \mathscr{A}_i)$, $i = 0, 1, \dots, n$ m.b. Räume, κ_i (sub-)stochastischer Kern von $\big(\times_{j=0}^{i-1} S_j, \bigotimes_{j=0}^{i-1} \mathscr{A}_j\big)$ nach $(S_i, \mathscr{A}_i)$ (oder von $(S_{i-1}, \mathscr{A}_{i-1})$ nach $(S_i, \mathscr{A}_i)$) für $i = 1, \dots, n$.

Setze (rekursiv)

$$\bigotimes_{j=1}^{i} \kappa_i = \kappa_1 \otimes \cdots \otimes \kappa_i := \big(\kappa_1 \otimes \cdots \otimes \kappa_{i-1}\big) \otimes \kappa_i.$$

$\bigotimes_{j=1}^{i} \kappa_i$ ist substochastischer Kern von $(S_0, \mathscr{A}_0)$ nach $\big(\times_{j=1}^{i} S_j, \bigotimes_{j=1}^{i} \mathscr{A}_j\big)$ (stochastisch, falls dies für alle κ_j gilt).

Ist weiter μ endliches Maß auf $(S_0, \mathscr{A}_0)$, so ist $\mu_i := \mu \otimes \bigotimes_{j=1}^{i} \kappa_i$ endliches Maß auf $\big(\times_{j=0}^{i} S_j, \bigotimes_{j=0}^{i} \mathscr{A}_j\big)$; $\mu \otimes \bigotimes_{j=1}^{i} \kappa_i$ ist W'maß, wenn $\mu(S_0) = 1$ und alle κ_j stochastisch sind.

(Verwende induktiv Satz 5.6 und Kor. 5.7.)

Wir betrachten folgende Situation:

$$\begin{aligned}&(\Omega_i, \mathscr{A}_i),\ i = 0, 1, 2, \ldots \quad \text{messbare Räume},\\ &\Omega^{(i)} := \bigtimes_{j=0}^{i} \Omega_j \ \text{ versehen mit } \ \mathscr{A}^{(i)} := \bigotimes_{j=0}^{i} \mathscr{A}_j,\\ &\Omega := \bigtimes_{j=0}^{\infty} \Omega_j \ \text{ versehen mit } \ \mathscr{A} := \bigotimes_{j=0}^{\infty} \mathscr{A}_j.\end{aligned}$$

Weiter sei

P_0 ein W'maß auf $(\Omega_0, \mathscr{A}_0)$,
κ_i stochastischer Kern von $(\Omega^{(i-1)}, \mathscr{A}^{(i-1)})$ nach $(\Omega_i, \mathscr{A}_i)$ für $i = 0, 1, 2, \ldots$,
$P_i := P_0 \otimes \bigotimes_{j=1}^{i} \kappa_j$ (ist ein W'maß auf $(\Omega^{(i)}, \mathscr{A}^{(i)})$ nach Kor. 5.8)

Beobachtung. Für $i \in \mathbb{N}$, $A \in \mathscr{A}^{(i)}$ gilt

$$P_{i+1}(A \times \Omega_{i+1}) = P_i(A)$$

Sei

$$\mathscr{C} := \left\{ A \times \Omega_{i+1} \times \Omega_{i+2} \times \cdots : i \in \mathbb{N},\ A \in \mathscr{A}^{(i)} \right\}$$

$\mathscr{C}$ *ist* eine Algebra mit $\sigma(\mathscr{C}) = \mathscr{A}$.

Satz 5.9 (Ionescu-Tulcea[1]). *Es gibt genau ein W'maß P auf $(\Omega, \mathscr{A})$, das*

$$P\Big(A \times \bigtimes_{j=k+1}^{\infty} \Omega_j\Big) = P_k(A) \quad \textit{für } A \in \mathscr{A}^{(k)},\ k \in \mathbb{N} \tag{5.4}$$

erfüllt.

Beweis. Eindeutigkeit ✓

Definiere P auf $\mathscr{C}$ via (5.4) : *ist* wohldefiniert, *ist* Inhalt (mit $P(\Omega) = 1$)).

Zeige: Seien $A_1 \supset A_2 \supset \cdots$, $A_n \in \mathscr{C}$ mit $\alpha := \inf_{n \in \mathbb{N}} P(A_n) > 0$, so gilt $\bigcap_n A_n \neq \varnothing$

O.E. sei $A_n = A'_n \times \Omega_{n+1} \times \Omega_{n+2} \times \cdots$ mit einem $A'_n \in \mathscr{A}^{(n)}$, setze

$$h_{m,n}(\omega_0, \ldots, \omega_m) := \Big(\bigotimes_{j=m+1}^{n} \kappa_j \Big)\big((\omega_0, \ldots, \omega_m), A'_n\big) \quad \text{für } n \geq m,\ \omega_i \in \Omega_i.$$

Für $n = m$ lese das triviale Produkt als „Identitätskern“, d.h.

$$h_{m,m}(\omega_0, \ldots, \omega_m) = \mathbf{1}_{A'_m}\big((\omega_0, \ldots, \omega_m)\big)$$

[1]Cassius Ionescu-Tulcea, *1923

Es gilt

$$\begin{aligned}
h_{m,n+1}(\omega_0,\dots,\omega_m) &= \Big(\bigotimes_{j=m+1}^{n+1}\kappa_j\Big)\big((\omega_0,\dots,\omega_m),A'_{n+1}\big)\\
&\le \Big(\bigotimes_{j=m+1}^{n+1}\kappa_j\Big)\big((\omega_0,\dots,\omega_m),A'_n\times\Omega_{n+1}\big)\\
&= \Big(\bigotimes_{j=m+1}^{n}\kappa_j\Big)\big((\omega_0,\dots,\omega_m),A'_n\big) = h_{m,n}(\omega_0,\dots,\omega_m),
\end{aligned}$$

also

$$h_{m,n}(\omega_0,\dots,\omega_m) \underset{n\to\infty}{\searrow} \inf_{n\ge m} h_{m,n}(\omega_0,\dots,\omega_m) =: h_m(\omega_0,\dots,\omega_m)$$

und

$$\int\limits_{\Omega_0\times\dots\times\Omega_m} h_m\,dP_m = \inf_{n\ge m}\int\limits_{\Omega_0\times\dots\times\Omega_m} h_{m,n}\,dP_m = \inf_{n\ge m} P_n(A'_n) = \inf_{n\ge m} P(A_n) > 0.$$

Zeige:

$$\text{es gibt } (\widetilde\omega_0,\widetilde\omega_1,\widetilde\omega_2,\dots)\in\Omega \ \text{ mit } \ h_m(\widetilde\omega_0,\widetilde\omega_1,\dots,\widetilde\omega_m)\ge\alpha\ \forall\, m \tag{5.5}$$

induktiv nach m:

$m = 0$: ✓

Seien $\widetilde\omega_0,\widetilde\omega_1,\dots,\widetilde\omega_m$ gewählt, so dass (5.5) (für dieses m) gilt:

$$\begin{aligned}
&\int_{\Omega_{m+1}} h_{m+1}(\widetilde\omega_0,\dots,\widetilde\omega_m,\omega_{m+1})\,\kappa_{m+1}\big((\widetilde\omega_0,\dots,\widetilde\omega_m),d\omega_{m+1}\big)\\
&= \inf_{n\ge m+1}\int_{\Omega_{m+1}} h_{m+1,n}(\widetilde\omega_0,\dots,\widetilde\omega_m,\omega_{m+1})\,\kappa_{m+1}\big((\widetilde\omega_0,\dots,\widetilde\omega_m),d\omega_{m+1}\big)\\
&= \inf_{n\ge m+1} h_{m,n}(\widetilde\omega_0,\dots,\widetilde\omega_m) = h_m(\widetilde\omega_0,\dots,\widetilde\omega_m)\ge\alpha
\end{aligned}$$

Nach Konstruktion ist

$$\alpha\le h_{m,m}(\widetilde\omega_0,\widetilde\omega_1,\dots,\widetilde\omega_m) = \mathbf{1}_{A'_m}\big((\widetilde\omega_0,\widetilde\omega_1,\dots,\widetilde\omega_m)\big) \quad \text{für alle } m\in\mathbb{N},$$

somit

$$(\widetilde\omega_0,\widetilde\omega_1,\dots,)\in\bigcap_{m=1}^{\infty}A_m\ne\varnothing.$$

□

Korollar 5.10 (Abzählbar unendliches Produktmaß)**.** *Seien $(\Omega_i,\mathscr{A}_i,P_i)$, $i\in\mathbb{N}_0$ W'räume. Es gibt genau ein W'maß P, geschrieben $P=\bigotimes_{i=0}^{\infty}P_i$, auf $\Omega := \times_{i=0}^{\infty}\Omega_i$, $\mathscr{A} := \bigotimes_{i=0}^{\infty}\mathscr{A}_i$ mit*

$$P\big(A_0\times\dots\times A_n\times\Omega_{n+1}\times\cdots\big) = \prod_{i=0}^{n}P_i(A_i) \quad \textit{für } n\in\mathbb{N},\ A_i\in\mathscr{A}_i.$$

(Wähle $\kappa_i\big((\omega_0,\dots,\omega_{i-1}),\cdot\big) = P_i(\cdot)$ in Satz 5.9.)

Korollar 5.11. *S endl. oder abzählbar, μ W'maß auf S, $(p_{x,y})_{x,y\in S}$ stochastische Matrix auf S. Es gibt genau ein W'maß P auf $S^{\mathbb{Z}_+}$ mit*

$$P\big(\{x_0\}\times\{x_1\}\times\cdots\{x_n\}\times S^{\mathbb{N}}\big) = \mu_{x_0}p_{x_0,x_1}\cdots p_{x_{n-1},x_n} \quad \textit{für } n\in\mathbb{N}, x_0,\dots,x_n\in S$$

5.2 Faltung

Definition 5.12. Für messbare $f, g : \mathbb{R}^d \to [0, \infty]$ ist die *Faltung* $f * g : \mathbb{R}^d \to [0, \infty]$ gegeben durch

$$(f * g)(x) = \int_{R^d} f(y)g(x-y)\,\lambda^d(dy), \quad x \in \mathbb{R}^d$$

Für endliche Maße μ, ν auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ ist die Faltung $\mu * \nu$ gegeben durch

$$(\mu * \nu)(A) = \int_{\mathbb{R}^d} \nu(A - x)\,\mu(dx), \quad A \in \mathcal{B}(\mathbb{R}^d)$$

Lemma 5.13. *$f * g : \mathbb{R}^d \to [0, \infty]$ ist messbar, es gelten $f * g = g * f$ und $\int (f * g)\,d\lambda = \left(\int f\,d\lambda\right)\left(\int g\,d\lambda\right)$.*

$\mu * \nu = \nu * \mu$ ist ein endliches Maß mit $(\mu * \nu)(\mathbb{R}^d) = \mu(\mathbb{R}^d)\nu(\mathbb{R}^d)$.

Beweis. Dies folgt aus Satz 5.2. □

Satz 5.14. *i) X, Y unabhängige $\mathbb{R}^d$-wertige ZVn mit Dichte f_X bzw. f_Y, so hat $X + Y$ die Dichte $f_X * f_Y$.*

ii) $\mu = f\,\lambda^d$, $\nu = g\,\lambda^d$ endliche Maße auf $\mathbb{R}^d$ mit Dichten bzgl. dem d-dimensionalen Lebesgue-Maß λ^d, so ist $\mu * \nu = (f * g)\,\lambda^d$

Beweis. i) Sei $x \in \mathbb{R}^d$, $A = \{(u, v) \in \mathbb{R}^d \times \mathbb{R}^d : u + v \le x\}$:

$$\begin{aligned}
\mathbb{P}(X + Y \le x) &= \mathbb{P}\big((X, Y) \in A\big) \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbf{1}_A(u, v) f_X(u) f_Y(v)\,(\lambda^d)^{\otimes 2}\big(d(u, v)\big) \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \mathbf{1}_A(u, v) f_X(u)\,\lambda^d(du) \right) f_Y(v)\,\lambda^d(dv) \\
&= \int_{\mathbb{R}^d} \left(\int_{(-\infty, x-v]} f_X(u)\,\lambda^d(du) \right) f_Y(v)\,\lambda^d(dv) \\
&= \int_{\mathbb{R}^d} \left(\int_{(-\infty, x]} f_X(u - v)\,\lambda^d(du) \right) f_Y(v)\,\lambda^d(dv) \\
&= \int_{(-\infty, x]} \left(\int_{\mathbb{R}^d} f_X(u - v) f_Y(v)\,\lambda^d(dv) \right) \lambda^d(du) = \int_{(-\infty, x]} (f_X * f_Y)\,d\lambda^d
\end{aligned}$$

ii) Analoge Rechnung □

Kapitel 6

Zur bedingten Erwartung

Erinnerung 6.1. Sei $(\Omega, \mathscr{F}, \mathbb{P})$ W'raum, $B \in \mathscr{F}$ mit $\mathbb{P}(B) > 0$.

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad A \in \mathscr{F}$$

heißt die *bedingte Wahrscheinlichkeit von A, gegeben B.*

$\mathbb{P}(\cdot \mid B)$ definiert durch diese Formel ein W'Maß auf $(\Omega, \mathscr{F})$ mit $\mathbb{P}(B \mid B) = 1$, für $Y \in \mathcal{L}^1(\mathbb{P})$ ist

$$\mathbb{E}[Y \mid B] = \int Y(\omega)\, \mathbb{P}(d\omega \mid B) = \frac{\mathbb{E}[\mathbf{1}_B Y]}{\mathbb{P}(B)}.$$

Bemerkung 6.2 (Diskreter Fall der bedingten Erwartung)**.** X und Y ZVn (auf einem W'raum $(\Omega, \mathscr{F}, \mathbb{P})$), $Y \in \mathcal{L}^1(\mathbb{P})$, X nehme höchstens abzählbar viele Werte $x_1, x_2, \dots$ an. Setze

$$f(x) := \mathbb{E}[Y \mid \{X = x\}] \quad \text{für } x \in \{x_1, x_2, \dots\}$$

und $\mathbb{E}[Y \mid X] := f(X)$.

Offenbar ist $\mathbb{E}[Y \mid X]$ $\sigma(X)$-messbar (es ist eine Funktion von X) und für $A \in \sigma(X)$ gilt

$$\mathbb{E}[Y \mathbf{1}_A] = \mathbb{E}\big[\mathbb{E}[Y \mid X] \mathbf{1}_A\big]$$

(klar für $A = \{X = x_i\}$ und dann auch für $A = \{X \in B\}$ mit $B \subset \{x_1, x_2, \dots\}$; jedes $A \in \sigma(X)$ hat diese Form).

Wir betrachten im folgenden einen W'raum $(\Omega, \mathscr{F}, \mathbb{P})$.

Definition 6.3. $X \in \mathcal{L}^1(\mathbb{P})$, $\mathscr{G} \subset \mathscr{F}$ Teil-σ-Algebra. Eine reellwertige ZV Y heißt (eine Version der) bedingte(n) Erwartung von X gegeben $\mathscr{G}$ (geschrieben $\mathbb{E}[X \mid \mathscr{G}]$), wenn gilt

i) Y ist $\mathscr{G}$-messbar,

ii) $\mathbb{E}[Y \mathbf{1}_A] = \mathbb{E}[X \mathbf{1}_A]$ für alle $A \in \mathscr{G}$.

Bemerkung. Äquivalent kann ii) durch ii') ersetzt werden:

ii') $\mathbb{E}[Y \cdot H] = \mathbb{E}[X \cdot H]$ für alle reellwertigen, beschr., $\mathcal{G}$-messbaren ZV H.

Falls $\mathcal{G} = \sigma(Z)$ für eine Zufallsvariable Z, so schreibt man oft auch $\mathbb{E}[X \mid Z] := \mathbb{E}[X \mid \sigma(Z)]$.

Satz 6.4. *Für $X \in \mathscr{L}^1(\mathbb{P})$ existiert $\mathbb{E}[X \mid \mathcal{G}]$ und ist eindeutig (bis auf $\mathbb{P}$-f.s.-Gleichheit).*

Beweis von Satz 6.4 (Eindeutigkeit). Seien $Y, \widetilde{Y}$ bedingte Erwartungen.

$$\mathbb{E}[Y \cdot \mathbf{1}_{\{Y > \widetilde{Y}\}}] = \mathbb{E}[X \cdot 1_{\{Y > \widetilde{Y}\}}] = \mathbb{E}[\widetilde{Y} \cdot \mathbf{1}_{\{Y > \widetilde{Y}\}}],$$

also $\mathbb{E}\big[(Y - \widetilde{Y}) \cdot \mathbf{1}_{\{Y > \widetilde{Y}\}}\big] = 0 \Rightarrow \mathbb{P}(Y > \widetilde{Y}) = 0$; analog gilt $\mathbb{P}(\widetilde{Y} > Y) = 0$, also $Y = \widetilde{Y}$ $\mathbb{P}$-f.s. □

Satz 6.5. *$\mathscr{H} \subset \mathcal{L}^2(\mathbb{P})$ abgeschlossener Unterraum. Zu $X \in \mathcal{L}^2(\mathbb{P})$ gibt es (bis auf $\mathbb{P}$-f.s. Gleichheit) genau ein $Y \in \mathscr{H}$ mit*

i) $\|Y - X\|_2 = \inf\big\{\|W - X\|_2 : W \in \mathscr{H}\big\}$

ii) $\mathbb{E}\big[(X - Y)Z\big] = 0$ *für alle* $Z \in \mathscr{H}$

Y ist (die Äquivalenzklasse) der orthogonalen Projektion von X auf $\mathscr{H}$, auch $\mathrm{Proj}_{\mathscr{H}}(X)$ *geschrieben.*

Beweis. Wähle $Y_n \in \mathscr{H}$ mit

$$\|X - Y_n\|_2 \underset{n\to\infty}{\longrightarrow} \alpha := \inf\big\{\|W - X\|_2 : W \in \mathscr{H}\big\}$$

Es ist

$$\|X - Y_n\|_2^2 + \|X - Y_m\|_2^2 = 2\left\|X - \frac{1}{2}(Y_n + Y_m)\right\|^2 + 2\left\|\frac{1}{2}(Y_n - Y_m)\right\|^2$$

wegen $\lim\limits_{n\to\infty} \|X - Y_n\|_2^2 = \lim\limits_{m\to\infty} \|X - Y_m\|_2^2 = \alpha^2 \le \liminf\limits_{n,m\to\infty} \left\|X - \frac{1}{2}(Y_n + Y_m)\right\|^2$ folgt

$$\limsup_{n,m\to\infty} \|Y_n - Y_m\|_2 = 0,$$

d.h. $(Y_n)_n \subset \mathcal{L}^2(\mathbb{P})$ ist Cauchy-Folge.

Demnach gibt es $Y \in \mathcal{L}^2(\mathbb{P})$ mit $Y_n \overset{\mathcal{L}^2(\mathbb{P})}{\underset{n\to\infty}{\longrightarrow}} Y$ ($\mathcal{L}^2(\mathbb{P})$ ist vollständig, Satz 3.25). Wähle (mit Bem. 3.23 und Lemma 3.24) Teilfolge $(n_k)_k$ mit $Y = \lim\limits_{k\to\infty} Y_{n_k}$ $\mathbb{P}$-f.s., dann ist $Y \in \mathscr{H}$ mit

$$\alpha^2 \le \mathbb{E}\big[(X - Y)^2\big] \le \liminf_{n\to\infty} \mathbb{E}\big[(X - Y_n)^2\big] = \alpha^2$$

Sei $Z \in \mathscr{H}$, für $t \in \mathbb{R}$ ist

$$\mathbb{E}\big[(X - (Y - tZ))^2\big] \ge \mathbb{E}\big[(X - Y)^2\big]$$

(nach Wahl von Y), also

$$2t\,\mathbb{E}\big[(X-Y)Z\big]+t^2\mathbb{E}[Z^2]\ge 0 \quad \text{für alle } t\in\mathbb{R},$$

was $\mathbb{E}\big[(X-Y)Z\big]=0$ erzwingt.

(Übrigens: Für $Y\in\mathscr{H}$ gilt *i)* $\Leftrightarrow$ *ii)*.)

Zur Eindeutigkeit: $\widetilde{Y}$ erfülle ebenfalls *i)*, es ist

$$\begin{aligned}\Big(X-\frac{1}{2}(Y+\widetilde{Y})\Big)^2 &= \Big(\frac{1}{2}(X-Y)+\frac{1}{2}(X-\widetilde{Y})\Big)^2\\ &\le \frac{1}{2}(X-Y)^2+\frac{1}{2}(X-\widetilde{Y})^2,\end{aligned}$$

die Ungleichung ist strikt auf $\{Y\ne\widetilde{Y}\}$.

Wäre $\mathbb{P}(Y\ne\widetilde{Y})>0$, so wäre

$$\mathbb{E}\Big[\Big(X-\frac{1}{2}(Y+\widetilde{Y})\Big)^2\Big]<\frac{1}{2}\mathbb{E}\big[(X-Y)^2\big]+\frac{1}{2}\mathbb{E}\big[(X-\widetilde{Y})^2\big]=\frac{1}{2}\alpha+\frac{1}{2}\alpha=\alpha,$$

ein Widerspruch, da $(Y+\widetilde{Y})/2\in\mathscr{H}$. □

Beweis von Satz 6.4 (Existenz). Sei zunächst $X\in\mathcal{L}^2(\mathbb{P})$,

$$\mathscr{H}:=\big\{Y\in\mathcal{L}^2(\mathbb{P}):Y \text{ ist } \mathscr{G}\text{-messbar}\big\}\subset\mathcal{L}^2(\mathbb{P})$$

ist ein abgeschlossener Unterraum, $Y:=\mathrm{Proj}_{\mathscr{H}}(X)$ (nach Satz 6.5) leistet das Gewünschte.

Beachte:

$$X\ge 0 \;\Rightarrow\; Y:=\mathbb{E}[X\mid\mathscr{G}]\ge 0 \quad \text{f.s.},$$

denn dann ist $0\le\mathbb{E}[X\mathbf{1}_{\{Y<0\}}]=\mathbb{E}[Y\mathbf{1}_{\{Y<0\}}]$.

Sei nun $X\in\mathcal{L}^1(\mathbb{P})$, $X\ge 0$:

$$Y_n:=\mathbb{E}\big[X\wedge n\big|\mathscr{G}\big]\underset{n\to\infty}{\nearrow} Y\ (\ge 0) \quad (\text{f.s.})$$

($X\wedge n\in\mathcal{L}^1(\mathbb{P})$, nutze dann obige Monotonie-Eigenschaft, um $\mathbb{E}[X\wedge n\mid\mathscr{G}]\le\mathbb{E}[X\wedge(n+1)\mid\mathscr{G}]$ f.s. zu sehen), für $A\in\mathscr{G}$ gilt

$$\mathbb{E}[X\mathbf{1}_A]=\lim_{n\to\infty}\mathbb{E}\big[(X\wedge n)\mathbf{1}_A\big]=\lim_{n\to\infty}\mathbb{E}[Y_n\mathbf{1}_A]=\mathbb{E}[Y\mathbf{1}_A]$$

(mit monotoner Konvergenz).

Für allgemeines $X\in\mathcal{L}^1(\mathbb{P})$ leistet

$$Y:=\mathbb{E}[X^+\mid\mathscr{G}]-\mathbb{E}[X^-\mid\mathscr{G}]$$

das Gewünschte. □

Satz 6.6. *$X, Y \in \mathcal{L}^1(\mathbb{P})$, $\mathscr{G} \subset \mathscr{F}$ Teil-σ-Algebra.*

i) $\mathbb{E}[aX + bY \mid \mathscr{G}] = a\mathbb{E}[X \mid \mathscr{G}] + b\mathbb{E}[Y \mid \mathscr{G}]$ *f.s. für $a, b \in \mathbb{R}$ (Linearität)*

ii) $X \leq Y$ *f.s.* $\Rightarrow$ $\mathbb{E}[X \mid \mathscr{G}] \leq \mathbb{E}[Y \mid \mathscr{G}]$ *f.s. (Monotonie)*

iii) $\big|\mathbb{E}[X \mid \mathscr{G}]\big| \leq \mathbb{E}\big[|X| \big| \mathscr{G}\big]$ *f.s. (Dreiecksungleichung)*

iv) Es gelte $\mathbb{E}\big[|XY|\big] < \infty$ und Y sei $\mathscr{G}$-messbar, dann ist

$$\mathbb{E}[XY \mid \mathscr{G}] = Y \cdot \mathbb{E}[X \mid \mathscr{G}] \text{ f.s.,}$$

insbesondere ist $\mathbb{E}[Y|\mathscr{G}] = Y$ f.s. („Herausziehen von Bekanntem“)

v) Sind $\sigma(X)$ und $\mathscr{G}$ unabhängig, so ist $\mathbb{E}[X \mid \mathscr{G}] = \mathbb{E}[X]$ f.s. (Verhalten bei Unabhängigkeit)

vi) $\mathscr{G}' \subset \mathscr{G}$ Teil-σ-Algebra, so ist

$$\mathbb{E}\big[\mathbb{E}[X \mid \mathscr{G}] \big| \mathscr{G}'\big] = \mathbb{E}[X \mid \mathscr{G}'] \text{ f.s.,}$$

(„Turmeigenschaft“) insbesondere ist

$$\mathbb{E}[\mathbb{E}[X \mid \mathscr{G}]] = \mathbb{E}[X]$$

vii) $0 \leq X_n \nearrow X$ f.s. für $n \to \infty$, so gilt

$$\mathbb{E}[X_n \mid \mathscr{G}] \underset{n\to\infty}{\nearrow} \mathbb{E}[X \mid \mathscr{G}] \quad \text{f.s. und in } \mathcal{L}^1(\mathbb{P})$$

(monotone Konvergenz)

viii) X_n reelle ZVn mit $|X_n| \leq Y$ $\forall n$ und $X_n \to X$ f.s., so gilt

$$\mathbb{E}[X_n \mid \mathscr{G}] \xrightarrow[n\to\infty]{} \mathbb{E}[X \mid \mathscr{G}] \quad \text{f.s. und in } \mathcal{L}^1(\mathbb{P})$$

(dominierte Konvergenz)

Beweis. (Die folgenden Gleichungen, etc. gelten jeweils f.s., auch wenn wir dies in der Notation nicht explizit machen.)
i) $a\mathbb{E}[X \mid \mathscr{G}] + b\mathbb{E}[Y \mid \mathscr{G}]$ ist $\mathscr{G}$-messb., und für $A \in \mathscr{G}$ gilt

$$\begin{aligned}\mathbb{E}\Big[\mathbf{1}_A\big(a\mathbb{E}[X \mid \mathscr{G}] + b\mathbb{E}[Y \mid \mathscr{G}]\big)\Big] &= a\mathbb{E}\big[\mathbf{1}_A\mathbb{E}[X \mid \mathscr{G}]\big] + b\mathbb{E}\big[\mathbf{1}_A\mathbb{E}[Y \mid \mathscr{G}]\big] \\ &= a\mathbb{E}[\mathbf{1}_A X] + b\mathbb{E}[\mathbf{1}_A Y] = \mathbb{E}\big[\mathbf{1}_A(aX + bY)\big]\end{aligned}$$

ii) $A := \big\{\mathbb{E}[X \mid \mathscr{G}] > \mathbb{E}[Y \mid \mathscr{G}]\big\} \in \mathscr{G}$ und

$$0 \leq \mathbb{E}\big[\mathbf{1}_A \underbrace{(Y - X)}_{\geq 0 \text{ f.s.}}\big] = \mathbb{E}\Big[\mathbf{1}_A\big(\mathbb{E}[Y \mid \mathscr{G}] - \mathbb{E}[X \mid \mathscr{G}]\big)\Big] \leq 0,$$

also $\mathbb{P}(A) = 0$.

iii) $\left|\mathbb{E}[X \mid \mathscr{G}]\right| \overset{i)}{=} \left|\mathbb{E}[X^+ \mid \mathscr{G}] - \mathbb{E}[X^- \mid \mathscr{G}]\right| \le \mathbb{E}[X^+ \mid \mathscr{G}] + \mathbb{E}[X^- \mid \mathscr{G}] \overset{i)}{=} \mathbb{E}\big[|X| \,\big|\, \mathscr{G}\big]$

iv) Seien zunächst $X, Y \ge 0$, $Y_n := 2^{-n}\lfloor 2^n Y \rfloor \wedge n$ $(\nearrow Y)$.

Für $A \in \mathscr{G}$ ist

$$\begin{aligned}\mathbb{E}\big[\mathbf{1}_A Y \mathbb{E}[X \mid \mathscr{G}]\big] &= \lim_{n\to\infty} \mathbb{E}\big[\mathbf{1}_A Y_n \mathbb{E}[X \mid \mathscr{G}]\big] \\ &= \lim_{n\to\infty} \mathbb{E}\Big[\mathbf{1}_A \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbf{1}_{\{Y = k/2^n\}} \mathbb{E}[X \mid \mathscr{G}]\Big] \\ &= \lim_{n\to\infty} \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbb{E}\big[\mathbf{1}_{A\cap\{Y=k/2^n\}} X\big] \\ &= \lim_{n\to\infty} \mathbb{E}\big[\mathbf{1}_A Y_n X\big] = \mathbb{E}[\mathbf{1}_A Y X]\end{aligned}$$

Für den allg. Fall zerlege $X = X^+ - X^-$, $Y = Y^+ - Y^-$, verwende *i)*.

v) Für $A \in \mathscr{G}$ gilt $\mathbb{E}[\mathbf{1}_A X] = \mathbb{E}[\mathbf{1}_A]\mathbb{E}[X] = \mathbb{E}\big[\mathbf{1}_A \mathbb{E}[X]\big]$

vi) Sei $A \in \mathscr{G}'$

$$\mathbb{E}\big[\mathbf{1}_A \mathbb{E}[X \mid \mathscr{G}]\big] = \mathbb{E}[\mathbf{1}_A X] = \mathbb{E}\big[\mathbf{1}_A \mathbb{E}[X \mid \mathscr{G}']\big]$$

d.h. $\mathbb{E}[X \mid \mathscr{G}']$ erfüllt die Def. von $\mathbb{E}\big[\mathbb{E}[X \mid \mathscr{G}] \,\big|\, \mathscr{G}'\big]$.

Zum Zusatz: Stets ist $\mathbb{E}[X \mid \{\varnothing, \Omega\}] = \mathbb{E}[X]$

vii) $\mathbb{E}[X_n \mid \mathscr{G}]$ ist (f.s.) nicht-fallend in n nach *ii)*,

$$\mathbb{E}\Big[\big|\mathbb{E}[X \mid \mathscr{G}] - \mathbb{E}[X_n \mid \mathscr{G}]\big|\Big] \overset{iii)}{\le} \mathbb{E}\Big[\mathbb{E}\big[|X - X_n| \,\big|\, \mathscr{G}\big]\Big] \overset{iv)}{=} \mathbb{E}\big[|X - X_n|\big] \underset{n\to\infty}{\longrightarrow} 0$$

also $\mathbb{E}[X_n \mid \mathscr{G}] \to \mathbb{E}[X \mid \mathscr{G}]$ in $\mathcal{L}^1(\mathbb{P})$.

Weiterhin: Für monotone Folgen von ZVn impliziert $\mathcal{L}^1(\mathbb{P})$-Konvergenz f.s.-Konvergenz: Seien $Z_n \nearrow Z$ und $Z \to Z_n$ in $\mathcal{L}^1(\mathbb{P})$, dann auch $Z_n \nearrow Z = \sup_m Z_m$ stochastisch, $\mathbb{P}\Big(\bigcap_{m\ge n} \big\{|Z_m - Z| < \varepsilon\big\}\Big) = \mathbb{P}(Z - \varepsilon < Z_n) \longrightarrow_{n\to\infty} 1$, somit

$$\mathbb{P}\Big(\bigcup_n \bigcap_{m\ge n} \big\{|Z_m - Z| < \varepsilon\big\}\Big) = 1 \quad \text{für jedes } \varepsilon > 0.$$

viii) $0 \le Z_n := \sup_{m\ge n} |X_m - X|$ $(\le 2Y)$, $Z_n \searrow 0$ f.s. für $n \to \infty$.

$$\mathbb{E}\Big[\big|\mathbb{E}[X \mid \mathscr{G}] - \mathbb{E}[X_n \mid \mathscr{G}]\big|\Big] \overset{iii)}{\le} \mathbb{E}\Big[\mathbb{E}\big[|X - X_n| \,\big|\, \mathscr{G}\big]\Big] \overset{iv)}{=} \mathbb{E}\big[|X - X_n|\big] \le \mathbb{E}[Z_n] \underset{n\to\infty}{\longrightarrow} 0,$$

also $\mathbb{E}[X_n \mid \mathscr{G}] \to \mathbb{E}[X \mid \mathscr{G}]$ in $\mathcal{L}^1(\mathbb{P})$. Weiter ist

$$\mathbb{E}\Big[\lim_{n\to\infty} \mathbb{E}[Z_n \mid \mathscr{G}]\Big] \le \liminf_{n\to\infty} \mathbb{E}[Z_n] = 0$$

d.h. $\mathbb{E}[Z_n \mid \mathscr{G}] \to 0$ f.s., somit

$$\big|\mathbb{E}[X \mid \mathscr{G}] - \mathbb{E}[X_n \mid \mathscr{G}]\big| \le \mathbb{E}\big[|X - X_n| \,\big|\, \mathscr{G}\big] \le \mathbb{E}[Z_n \mid \mathscr{G}]\Big] \underset{n\to\infty}{\longrightarrow} 0 \quad \text{f.s.}$$

$\square$

Satz 6.7 (Jensen'sche Ungleichung für die bedingte Erwartung)**.** *$X \in \mathcal{L}^1(\mathbb{P})$, $k : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ konvex, dann gilt*

$$\mathbb{E}\big[k(X)\,\big|\,\mathscr{G}\big] \geq k\big(\mathbb{E}[X \mid \mathscr{G}]\big) \quad \textit{f.s.}$$

Beweis. Die Behauptung ist offenbar erfüllt (mit Satz 6.6, *i)*), wenn k affin-linear ist, d.h. $k(x) = ax + b$ mit gewissen $a, b \in \mathbb{R}$.

Sei nun k konvex und nicht von dieser Form. Man kann die Funktion k als das Supremum ihrer Stützgeraden schreiben, d.h.

$$k(x) = \sup\{ax + b : (a,b) \in S\} \quad \text{mit } S := \{(a,b) \in \mathbb{R}^2 : ay + b \leq k(y)\ \forall\, y \in \mathbb{R}\}$$

und es gilt auch ($(a,b) \mapsto ax + b$ ist stetig für jedes $x \in \mathbb{R}$ und da k keine Gerade ist, kann man zu $(a,b) \in S$ gewisse $(a_m, b_m) \in S \cap \mathbb{Q}^2$ finden mit $a_m \to a, b_m \to b$ für $m \to \infty$)

$$k(x) = \sup\{ax + b : (a,b) \in S \cap \mathbb{Q}^2\}.$$

Sei $(a_n, b_n)_{n \in \mathbb{N}}$ eine Aufzählung von $S \cap \mathbb{Q}^2$, für jedes n gilt

$$\mathbb{E}\big[k(X)\,\big|\,\mathscr{G}\big] \geq \mathbb{E}\big[a_n X + b_n\,\big|\,\mathscr{G}\big] = a_n \mathbb{E}[X \mid \mathscr{G}] + b_n \quad \text{f.s.}$$

(mit Satz 6.6, *ii)* und *i)*), daher auch (man muss oben nur abzählbar viele Ausnahmemengen betrachten)

$$\mathbb{E}\big[k(X)\,\big|\,\mathscr{G}\big] \geq \sup\big\{a_n \mathbb{E}[X \mid \mathscr{G}] + b_n : n \in \mathbb{N}\big\} = k\big(\mathbb{E}[X \mid \mathscr{G}]\big) \quad \text{f.s.}$$

□

Bemerkung 6.8. X, Y reelle ZVn mit gemeinsamer Dichte $f_{X,Y}$, d.h.

$$\mathbb{P}\big((X,Y) \in A\big) = \int_A f_{X,Y}(x,y)\,\lambda^{\otimes 2}\big(d(x,y)\big) \quad \text{für } A \in \mathcal{B}(\mathbb{R}^2),$$

$Y \in \mathcal{L}^1(\mathbb{P})$.

Sei für $x \in \mathbb{R}$

$$\begin{aligned} f_X(x) &:= \int_{\mathbb{R}} f_{X,Y}(x,y)\,\lambda(dy) \quad \text{die Marginaldichte von } X, \\ \varphi(x) &:= \int_{\mathbb{R}} y \frac{f_{X,Y}(x,y)}{f_X(x)}\,\lambda(dy) \mathbf{1}_{f_X(x) > 0}. \end{aligned}$$

Dann gilt

$$\mathbb{E}[Y | \sigma(X)] = \varphi(X) \ \text{ f.s.}$$

denn für $B \in \mathcal{B}(\mathbb{R})$ ist

$$\begin{aligned} \mathbb{E}\big[\mathbf{1}_{\{X \in B\}} \varphi(X)\big] &= \int_{\mathbb{R}} \mathbf{1}_B(x) \varphi(x) f_X(x)\,\lambda(dx) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_B(x) y f_{X,Y}(x,y) \lambda(dy)\ \lambda(dx) \\ &= \int_{\mathbb{R}^2} \mathbf{1}_B(x) y f_{X,Y}(x,y)\,\lambda^{\otimes 2}\big(d(x,y)\big) = \mathbb{E}\big[\mathbf{1}_{\{X \in B\}} Y\big]. \end{aligned}$$

Bericht 6.9. (Zu regulären Versionen bedingter Verteilungen) Wenn $Y = 1_B$ für ein Ereignis $B \in \mathscr{A}$, so schreibt man gelegentlich auch $\mathbb{P}(B \mid \mathscr{G}) = \mathbb{E}[Y \mid \mathscr{G}]$. Man muss allerdings etwas vorsichtig sein bei der Interpretation von $\mathbb{P}(\cdot \mid \mathscr{G})$ als ein (zufälliges) Maß, da i.A. überabzählbar viele B in Frage kommen und damit die Kompatibilität der in der Definition der bedingten Erwartung implizit vorkommenden Nullmengen (vgl. Def. 6.3) wenigstens a priori unklar bleibt.

In „gutartigen“ Fällen ist eine konsistente Wahl möglich, das Stichwort dazu lautet „reguläre bedingte Verteilung einer Zufallsvariable“. Wir skizzieren hier knapp den reellwertigen Fall:

Sei X reellwertige ZV (auf einem W'raum $(\Omega, \mathscr{F}, \mathbb{P})$), $\mathscr{G} \subset \mathscr{F}$ Teil-σ-Algebra. Dann gibt es einen stochastischen Kern $\kappa_{X|\mathscr{G}}$ von $(\Omega, \mathscr{G})$ nach $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit

$$\kappa_{X|\mathscr{G}}(\omega, B) = \mathbb{E}\big[\mathbf{1}_{\{X \in B\}} \big| \mathscr{G}\big](\omega) \text{ f.s. für alle } B \in \mathcal{B}(\mathbb{R}),$$

d.h. (vgl. Def. 5.4)

für jedes $B \in \mathcal{B}(\mathbb{R})$ ist $\omega \mapsto \kappa_{X|\mathscr{G}}(\omega, B)$ eine Version von $\mathbb{E}\big[\mathbf{1}_{\{X \in B\}} \big| \mathscr{G}\big]$ und
für jedes ω ist $\kappa_{X|\mathscr{G}}(\omega, \cdot)$ ein W'maß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Die Hauptidee besteht darin, das gewünschte Maß $\kappa_{X|\mathscr{G}}(\omega, \cdot)$ auf $\mathbb{R}$ anhand seiner Verteilungsfunktion zu charakterisieren (vgl. Satz 1.27), die zielführende Beobachtung ist dann, dass eine Verteilungsfunktion (wegen der Monotonie) bereits durch ihre Werte an abzählbar vielen Stellen festgelegt ist. Man betrachtet also $B = (-\infty, r], r \in \mathbb{Q}$ und setzt

$$F_r := \mathbb{E}[\mathbf{1}_{(-\infty, r]}(X) \mid \mathscr{G}].$$

Dann gilt (mit den Eigenschaften der bedingten Erwartung aus Satz 6.6) wie gewünscht $\mathbb{P}$-f.s.: $F_r \le F_{r'}$, für $r < r', (r, r' \in \mathbb{Q})$, $\lim_{n\to\infty} F_{r+\frac{1}{n}} = F_r$ für $r \in \mathbb{Q}$, $\lim_{n\to\infty} F_n = 1$, $\lim_{n\to-\infty} F_n = 0$.

Wegen der Abzählbarkeit von $\mathbb{Q}$ gibt es $N \in \mathscr{F}$ mit $\mathbb{P}(N) = 0$, so dass obiges für $\omega \in \Omega \setminus N$ und alle $r, r' \in \mathbb{Q}$ gilt. Dann definiert

$$\widetilde{F}_s := \begin{cases} \inf\{F_r : r \ge s, r \in \mathbb{Q}\}, & \omega \in \Omega \setminus N, \\ \overline{F}_s, & \omega \in N, \end{cases}$$

wobei $\overline{F}$ irgendeine Verteilungsfunktion ist, die (zufällige) Verteilungsfunktion von $k_{X,\mathscr{G}}$. Details finden sich beispielsweise in [Kl, Kap. 8.3, speziell Satz 8.28].

Man kann dieses Argument relativ leicht erweitern auf die Situation, dass der Wertebereich $(E, \mathscr{B})$ von X ein sogenannter Standard-Borel-Raum ist (auch der Name Borel'scher Raum ist üblich), d.h. wenn es ein $A \in \mathcal{B}(\mathbb{R})$ und eine Bijektion $\phi : E \to A$ gibt, so dass ϕ und ϕ^{-1} jeweils messbar sind (dann sind $(E, \mathscr{B})$ und $(A, \mathcal{B}(A))$ isomorph als messbare Räume). Dann ist nämlich $X' := \phi \circ X$ eine reellwertige ZV, und die Argumentation oben greift (vgl. auch [Kl, Satz 8.36]).

Schließlich kann man zeigen, dass jeder separable und vollständige metrische Raum E, versehen mit seiner Borel-σ-Algebra, ein Standard-Borel-Raum ist (siehe z.B. L.C.G. Rogers, D. Williams, *Diffusions, Markov processes and martingales*, Bd. 1, Ch. II.82; L. Breiman, *Probability*, Appendix 7). Solche Wertebereiche heißen *polnische Räume*, sie spielen in der allgemeinen Theorie der Stochastik eine wichtige Rolle (beispielsweise sind $\mathbb{R}^d$ oder $C([0,1])$ mit Supremumsnorm polnisch).

Kapitel 7

Martingale (in diskreter Zeit)

Martingale[1] sind (u.a.) eine mathematische Formalisierung des Begriffs des fairen Spiels und der Vorstellung, dass man dabei nicht auf systematische Weise gewinnen kann.

Wir betrachten sozusagen zum „Appetit-Anregen“ folgendes Beispiel:

Beispiel 7.1. Betrachte eine faire Münzwurffolge, d.h. seien $W_1, W_2, \dots$ unabhängig und identisch uniform verteilt auf $\{K, Z\}$. Sei

$$R := \min\Big\{k \in \mathbb{N} : (W_k, W_{k+1}, W_{k+2}, W_{k+3}) = (Z, K, Z, K)\Big\}.$$

Zum Zeitpunkt $R+3$ können wir sehen, dass das Muster (Z, K, Z, K) zum ersten Mal „fertig“ ist ($R+3$ ist eine sog. Stoppzeit, siehe Def. 7.11 unten).

Um $\mathbb{E}[R]$ zu bestimmen stellen wir uns ein „faires Casino“ vor:

Setze vor dem i-ten Wurf x Euro, erhalte $2x$ Euro oder 0 Euro je nach Ausgang (und jeder mögliche Ausgang hat W'keit 1/2). Spieler i steigt in Runde i in das Spiel ein und setzt einen Euro auf Z. Falls er gewinnt, setzt er in Runde $i+1$ zwei Euro auf K. Gewinnt er wieder, setzt er in Runde $i+2$ vier Euro auf Z. Sollte er wieder gewinnen, setzt er in Runde $i+3$ acht Euro auf K. Gewinnt er auch dieses Spiel, hört er auf. Sei nun $X_{i,n}$ der Gewinn des i-ten Spielers nach der n-ten Runde und

$$X_n := \sum_{i=1}^{n} X_{i,n}$$

der Gesamtgewinn aller Spieler nach Runde n. Aufgrund der „Fairness“ sollte gelten

$$0 = \mathbb{E}[X_0] = \mathbb{E}[X_n] = \mathbb{E}[X_{R+3}]. \qquad (*)$$

(Wir werden die Theorie hinter $(*)$ entwickeln, siehe Korollar 7.19 unten.)

[1]Zur farbigen Geschichte des Begriffs Martingal siehe beispielsweise den Artikel von Roger Mansuy, The origins of the word “martingale”, Electronic Journal for History of Probability and Statistics, Vol. 5 no. 1, (2009), `http://www.jehps.net`.

Zum Zeitpunkt $R+3$ hat Spieler R einen Gewinn von 15 Euro, Spieler $R+2$ hat einen Gewinn von 3 Euro und die anderen $R+1$ Spieler, die bisher mitgespielt haben, haben einen Gewinn von -1 Euro, das heißt

$$X_{R+3} = 15 + 3 - (R+1).$$

$(*)$ liefert

$$0 = \mathbb{E}[X_{R+3}] = \mathbb{E}[R-17],$$

und damit $\mathbb{E}[R] = 17$.

7.1 Grundlegendes

Im Folgenden sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum.

Definition 7.2. Eine Familie $(\mathscr{F}_n)_{n=0,1,\ldots}$ von (Teil-) σ-Algebren mit

$$\mathscr{F}_0 \subset \mathscr{F}_1 \subset \mathscr{F}_2 \subset \ldots \subset \mathscr{F}$$

heißt *Filtration*. $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n=0,1,\ldots}, \mathbb{P})$ heißt filtrierter Wahrscheinlichkeitsraum.

Bemerkung 7.3.
i) Interpretation: $\mathscr{F}_n$ enthält diejenigen Ereignisse, die bis zum Zeitpunkt n entschieden sind.

ii) Ist $X = (X_n)_{n=0,1,\ldots}$ eine Familie von Zufallsvariablen (ein sog. stochastischer Prozess), so ist $\mathscr{F}_n = \sigma(X_1, \ldots X_n)$, $n \in \mathbb{N}_0$ eine Filtration (die von X erzeugte Filtration).

Definition 7.4. Es sei $X = (X_n)_n$ ein stochastischer Prozess und $(\mathscr{F}_n)_n$ eine Filtration. X heißt *adaptiert* (an $(\mathscr{F}_n)_n$), wenn X_n $\mathscr{F}_n$-messbar ist für alle $n \in \mathbb{N}_0$.

Definition 7.5. Es sei $X = (X_n)_n$ ein (reellwertiger) stochastischer Prozess und $(\mathscr{F}_n)_n$ eine Filtration. X heißt ein *Martingal* (bzgl. $(\mathscr{F}_n)_n$ unter $\mathbb{P}$), wenn gilt:

i) X ist adaptiert (an $(\mathscr{F}_n)_n$).

ii) $X_n \in \mathcal{L}^1(\mathbb{P})$ für $n \in \mathbb{N}_0$.

iii) $\mathbb{E}[X_{n+1} \mid \mathscr{F}_n] = X_n$ f.s. für jedes $n \in \mathbb{N}_0$.

Falls in iii) $\mathbb{E}[X_{n+1} \mid \mathscr{F}_n] \geq X_n$ gilt, so heißt X ein *Submartingal*. Falls $\mathbb{E}[X_{n+1} \mid \mathscr{F}_n] \leq X_n$ gilt, so heißt X ein *Supermartingal*.

Bemerkung 7.6. Induktiv folgt für ein Martingal X

$$\mathbb{E}[X_n \mid \mathscr{F}_m] = X_m \text{ f.s. für alle } 0 \leq m \leq n.$$

(bzw. „$\geq$“ für ein Sub- und „$\leq$“ für ein Supermartingal.

Beispiel 7.7. i) Seien $Y_1, Y_2, \dots$ unabhängige, reelle Zufallsvariablen mit $Y_n \in \mathcal{L}^1(\mathbb{P})$ und $\mathbb{E}[Y_n] = 0$ für alle $n \in \mathbb{N}$. Sei $S_0 := 0$ und $S_n := Y_1 + Y_2 + \ldots + Y_n = S_{n-1} + Y_n$ für $n \in \mathbb{N}$. Dann ist $(S_n)_n$ ein Martingal bezgl. $(\mathscr{F}_n)_n$ mit $\mathscr{F}_n = \sigma(S_1, \dots S_n) = \sigma(Y_1, \dots Y_n)$, $n \in \mathbb{N}$ (und $\mathscr{F}_0 = \{\varnothing, \Omega\}$), denn $S_n \in \mathcal{L}^1(\mathbb{P})$ als Summe von $\mathcal{L}^1(\mathbb{P})$-Variablen und es gilt

$$\mathbb{E}[S_{n+1} \mid \mathscr{F}_n] = \mathbb{E}[S_n + Y_{n+1} \mid \mathscr{F}_n] = \underbrace{\mathbb{E}[S_n \mid \mathscr{F}_n]}_{=S_n \text{ f.s.}} + \underbrace{\mathbb{E}[Y_{n+1} \mid \mathscr{F}_n]}_{=\mathbb{E}[Y_{n+1}]=0 \text{ f.s.}} = S_n \text{ f.s.}$$

ii) Pólyas Urne: Eine Urne enthalte anfangs $s > 0$ schwarze und $w > 0$ weiße Kugeln. Ziehe jeweils eine Kugel rein zufällig und lege sie zusammen mit einer neuen Kugel der selben Farbe zurück. Sei X_n die Anzahl weißer Kugeln nach n Zügen und $A_n := \frac{X_n}{s+w+n}$ der Anteil weißer Kugeln in der Urne. Dann ist $(A_n)_n$ ein Martingal bzgl. $\mathscr{F}_n = \sigma(A_0, \dots, A_n)$, denn auf $\{X_n = k\}$ gilt

$$\begin{aligned}\mathbb{E}[A_{n+1} \mid \mathscr{F}_n] &= \frac{k}{s+w+n} \cdot \frac{k+1}{s+w+n+1} + \frac{w+s+n-k}{s+w+n} \cdot \frac{k}{s+w+n+1} \\ &= \frac{k}{s+w+n} \cdot \underbrace{\frac{k+1+w+s+n-k}{w+s+n+1}}_{=1} = A_n.\end{aligned}$$

Definition 7.8. Sei $(\mathscr{F}_n)_n$ eine Filtration. Ein stochastischer Prozess $(C_n)_n$ heißt *previsibel* (bzgl. $(\mathscr{F}_n)_n$), auch vorhersagbar, wenn C_n $\mathscr{F}_{n-1}$-messbar ist für jedes $n \in \mathbb{N}$ (C_0 spielt hier keine Rolle).

Definition 7.9. Sei $(X_n)_n$ adaptiert und $(C_n)_n$ previsibel bzgl. $(\mathscr{F}_n)_n$. Setze

$$(C \bullet X)_0 := 0, \quad (C \bullet X)_n := \sum_{m=1}^{n} C_m (X_m - X_{m-1}), \quad n \in \mathbb{N}. \tag{7.1}$$

Der Prozess $C \bullet X = ((C \bullet X)_n)_{n \in \mathbb{N}_0}$ heißt *(diskretes) stochastisches Integral* von C bezüglich X. $C \bullet X$ ist (offenbar) adaptiert.

Spielinterpretation: $C \bullet X$ ist ein akkumulierter Gewinnprozess für einen Spieler, der in der m-ten Runde jeweils C_m-fachen Einsatz setzt.

Lemma 7.10. *Es sei $(X_n)_n$ ein Martingal und $(C_n)_n$ ein previsibler Prozess bzgl. $(\mathscr{F}_n)_n$. Es gelte mindestens eine der folgenden drei Bedingungen*

i) $(C_n)_n$ ist lokal beschränkt, d.h. es gibt Konstanten c_n mit $|C_n| \le c_n$ f.s. für alle $n \in \mathbb{N}$.

ii) $(X_n - X_{n-1})_n$ ist lokal beschränkt und $C_n \in \mathcal{L}^1(\mathbb{P})$ für alle $n \in \mathbb{N}$.

iii) $X_n, C_n \in \mathcal{L}^2(\mathbb{P})$ für alle $n \in \mathbb{N}_0$.

Dann ist $C \bullet X$ ein Martingal. Ist $C_n \ge 0$ für alle $n \in \mathbb{N}_0$ und X ein Sub- bzw. Supermartingal, so auch $C \bullet X$.

Beweis. i), ii) oder iii) garantieren, dass $C_m(X_m - X_{m-1}) \in \mathcal{L}^1(\mathbb{P})$, denn für iii) gilt mit der Cauchy-Schwarz-Ungleichung

$$\mathbb{E}[|C_m(X_m - X_{m-1})|] \le (\mathbb{E}[C_m^2])^{\frac{1}{2}}(\mathbb{E}[(X_m - X_{m-1})^2])^{\frac{1}{2}} < \infty.$$

Also gilt

$$\mathbb{E}[(C \bullet X)_{n+1} \mid \mathscr{F}_n] = \underbrace{\mathbb{E}[C_{n+1}(X_{n+1} - X_n) \mid \mathscr{F}_n]}_{=C_{n+1}\cdot \underbrace{\mathbb{E}[X_{n+1}-X_n|\mathscr{F}_n]}_{=0} = 0 \text{ f.s.}} + \underbrace{\mathbb{E}[(C \bullet X)_n \mid \mathscr{F}_n]}_{=(C\bullet X)_n \text{ f.s.}} = (C \bullet X)_n \text{ f.s.}$$

□

Definition 7.11. Sei $(\mathscr{F}_n)_n$ eine Filtration. Eine Zufallsvariable T mit Werten in $\mathbb{N}_0 \cup \{\infty\}$ heißt eine $((\mathscr{F}_n)_n$-)*Stoppzeit*, wenn $\{T \le n\} \in \mathscr{F}_n$ für alle $n \in \mathbb{N}_0$ gilt. Für eine Stoppzeit T *ist*

$$\mathscr{F}_T := \{A \in \mathscr{F} \mid A \cap \{T \le n\} \in \mathscr{F}_n \text{ für alle } n \in \mathbb{N}\}$$

eine σ-Algebra, sie heißt die *(σ-Algebra der) T-Vergangenheit.*

Interpretation: $\mathscr{F}_T$ enthält diejenigen Ereignisse, die sich zu dem (zufälligen) Zeitpunkt T entscheiden lassen.

Bemerkung 7.12. T ist genau dann eine Stoppzeit, wenn $\{T = n\} \in \mathscr{F}_n$ für alle $n \in \mathbb{N}_0$, denn $\{T = n\} = \{T \le n\} \cap \{T \le n-1\}^{\mathsf{c}}$.

Beispiel 7.13.
i) Jede Konstante t_0 ist eine Stoppzeit.

ii) Es sei $(X_n)_n$ ein adaptierter stochastischer Prozess mit Werten in $(E, \mathscr{B})$ und $K \in \mathscr{B}$. Dann ist

$$T := \inf\{n \in \mathbb{N}_0 \mid X_n \in K\}$$

(mit Interpretation $\inf \varnothing = \infty$) eine Stoppzeit, denn $\{T \le n\} = \bigcup_{m=0}^{n} \{X_m \in K\} \in \mathscr{F}_n$.

Bemerkung. $L := \sup\{n \in \mathbb{N}_0 \mid X_n \in K\}$ ist im Allgemeinen keine Stoppzeit.

Lemma 7.14. *Sind σ, τ Stoppzeiten, so sind auch $\sigma \wedge \tau$, $\sigma \vee \tau$ und $\sigma + \tau$ Stoppzeiten.*

Beweis. Sei $n \in \mathbb{N}_0$. Es gilt

$$\{\sigma \vee \tau \le n\} = \{\sigma \le n\} \cap \{\tau \le n\} \in \mathscr{F}_n, \quad \{\sigma \wedge \tau \le n\} = \{\sigma \le n\} \cup \{\tau \le n\} \in \mathscr{F}_n.$$

Also sind $\sigma \wedge \tau$ und $\sigma \vee \tau$ Stoppzeiten. Dann sind auch $\sigma \wedge n$ und $\tau \wedge n$ Stoppzeiten, also gilt insbesondere für $m \le n$: $\{\sigma \wedge n \le m\}, \{\tau \wedge n \le m\} \in \mathscr{F}_m \subset \mathscr{F}_n$. Dann sind

$$\sigma' := \sigma \wedge n + \mathbf{1}_{\{\sigma > n\}}, \quad \tau' := \tau \wedge n + \mathbf{1}_{\{\tau > n\}}$$

$\mathscr{F}_n$-messbar, also ist auch $\sigma' + \tau'$ $\mathscr{F}_n$-messbar. Somit gilt

$$\{\sigma + \tau \le n\} = \{\sigma' + \tau' \le n\} \in \mathscr{F}_n,$$

also ist auch $\sigma + \tau$ eine Stoppzeit. □

Bemerkung. $\sigma - \tau$ ist im Allgemeinen keine Stoppzeit.

Lemma 7.15. *Sind σ, τ Stoppzeiten mit $\sigma \leq \tau$, dann gilt $\mathscr{F}_\sigma \subset \mathscr{F}_\tau$.*

Beweis. Sei $A \in \mathscr{F}_\sigma$ und $n \in \mathbb{N}_0$. Da $\{\tau \leq n\} \subset \{\sigma \leq n\}$, gilt

$$A \cap \{\tau \leq n\} = \underbrace{(A \cap \{\sigma \leq n\})}_{\in \mathscr{F}_n} \cap \{\tau \leq n\} \in \mathscr{F}_n.$$

□

Beobachtung 7.16. Sei $(\mathscr{F}_n)_n$ eine Filtration, T eine Stoppzeit mit $T < \infty$ f.s. und $(X_n)_n$ ein adaptierter stochastischer Prozess mit Werten in $(E, \mathscr{B})$. Dann ist $X_T = X_{T(\omega)}(\omega)$ eine $\mathscr{F}_T$-messbare Zufallsvariable und $X^{(T)} = (X_n^{(T)})_{n \in \mathbb{N}_0} = (X_{T \wedge n})_{n \in \mathbb{N}_0}$ ein adaptierter stochastischer Prozess.

Beweis. Sei $B \in \mathscr{B}$. Dann ist

$$\{X_T \in B\} = \bigcup_{n=0}^{\infty} \underbrace{\{T = n, X_n \in B\}}_{\in \mathscr{F}_n \subset \mathscr{F}} \in \mathscr{F},$$

also ist X_T $\mathscr{F}$-messbar. Ebenso gilt

$$\{X_T \in B\} \cap \{T \leq n\} = \bigcup_{k=0}^{n} \underbrace{\{T = k, X_k \in B\}}_{\in \mathscr{F}_k} \subset \mathscr{F}_n,$$

also ist X_T $\mathscr{F}_T$-messbar.

Weiter ist $X_n^{(T)} = X_{T \wedge n}$ messbar bzgl. $\mathscr{F}_{T \wedge n} \subset \mathscr{F}_n$, d.h. $(X_n^{(T)})_n$ ist adaptiert. □

Bemerkung 7.17. $(X_n^{(T)})_n$ ist auch adaptiert an $\mathscr{F}^{(T)} := (\mathscr{F}_{T \wedge n})_n$.

Lemma 7.18. *Sei T eine Stoppzeit. Ist $(X_n)_n$ ein (Sub- / Super-) Martingal, so auch $(X_n^{(T)})_n$.*

Beweis. Sei $C_n := \mathbf{1}_{\{T \geq n\}}, n \in \mathbb{N}$. $(C_n)_n$ ist previsibel, denn $\{T \geq n\} = \{T \leq n-1\}^\mathsf{c} \in \mathscr{F}_{n-1}$. Schreibe

$$X_{T \wedge n} = X_0 + \sum_{m=1}^{T \wedge n} (X_m - X_{m-1}) = X_0 + \sum_{m=1}^{n} \mathbf{1}_{\{T \geq m\}} (X_m - X_{m-1}) = (C \bullet X)_n + X_0.$$

Damit folgt die Behauptung aus Lemma 7.10. □

Korollar 7.19. *Sei X ein Supermartingal und T eine Stoppzeit. Es gelte mindestens eine der folgenden Bedingungen*

i) T ist beschränkt.

ii) X ist beschränkt, d.h. $\sup_{n \in \mathbb{N}_0} |X_n| \leq c$ *f.s. für ein $c \in \mathbb{R}_+$, und $T < \infty$ f.s.*

iii) $\mathbb{E}[T] < \infty$ *und* $\sup_{n\in\mathbb{N}} |X_n - X_{n-1}| \le c$ *f.s. für ein* $c \in \mathbb{R}_+$.

Dann gilt $\mathbb{E}[X_T] \le \mathbb{E}[X_0]$. *(Im Falle eines Martingals gilt Gleichheit.)*

Beweis. Angenommen *i)* gilt, dann existiert ein $m \in \mathbb{N}$ mit $T \le m$. Nach Lemma 7.18 ist $(X_{T\wedge n})_n$ ein Supermartingal, also gilt

$$\mathbb{E}[X_0] = \mathbb{E}[X_{T\wedge 0}] \ge \mathbb{E}[X_{T\wedge m}] = \mathbb{E}[X_T].$$

Gilt *ii)*, so ist $\mathbb{E}[X_T] = \lim_{m\to\infty} \mathbb{E}[X_{T\wedge m}] \le \mathbb{E}[X_0]$ mit dem Satz von der dominierten Konvergenz (Satz 3.13).

Gilt *iii)*, dann folgt $T < \infty$ f.s. und $X_{T\wedge m} \xrightarrow[m\to\infty]{} X_T$ f.s.
Es gilt

$$\sup_{m\in\mathbb{N}_0} |X_{T\wedge m}| \le \sup_{m\in\mathbb{N}_0} \Big(|X_0| + (T\wedge m)\cdot c \Big) \le |X_0| + cT.$$

Da $|X_0| + cT \in \mathcal{L}^1(\mathbb{P})$, ist dies eine integrable Majorante für $X_{T\wedge m}$ und es folgt wiederum mit dem Satz von der dominierten Konvergenz (Satz 3.13)

$$\mathbb{E}[X_T] = \lim_{m\to\infty} \mathbb{E}[X_{T\wedge m}] \overset{i)}{\le} \mathbb{E}[X_0].$$

□

Nochmal zu Beispiel 7.1. An dieser Stelle haben wir genug Theorie entwickelt, um Beispiel 7.1 „rigoros" behandeln zu können:

R und *X* dort erfüllen die Bedingung *iii)* von Korollor 7.19: Es gilt $\mathbb{P}(R > 4k) \le \mathbb{P}\Big((W_{4j+1}, W_{4j+2}, W_{4j+3}, W_{4j+4}) \ne (Z,K,Z,K) \text{ für } = j = 0,1,\dots,k-1\Big) = \Big(1-(1/2)^4\Big)^4$, somit $\mathbb{E}[R] = \sum_{r=1}^{\infty} \mathbb{P}(R \ge r) < \infty$ und $|X_n - X_{n-1}| \le 1+2+4+8 = 15$.

7.2 Martingalkonvergenzsatz

Sei $(X_n)_n$ ein adaptierter, reellwertiger Prozess und $-\infty < a < b < \infty$.
Setze $C_1 := \mathbf{1}_{\{X_0<a\}}$ und für $n > 1$ rekursiv (siehe auch Abbildung 7.1)

$$C_n := \mathbf{1}_{\{C_{n-1}=1,\ X_{n-1}\le b\}} + \mathbf{1}_{\{C_{n-1}=0,\ X_{n-1}<a\}}.$$

$(C_n)_n$ ist previsibel. Sei weiter

$$U_n^{(a,b)} := \sum_{k=1}^{n} \mathbf{1}_{\{C_k=1,\ C_{k+1}=0\}}$$

die Anzahl der abgeschlossenen Aufkreuzungen von unter a nach über b bis zur Zeit n. $U_n^{(a,b)}$ ist $\mathscr{F}_n$-messbar. Setze $Y := C \bullet X$, so gilt

$$Y_n \ge (b-a)U_n^{(a,b)} - (X_n - a)^-.$$

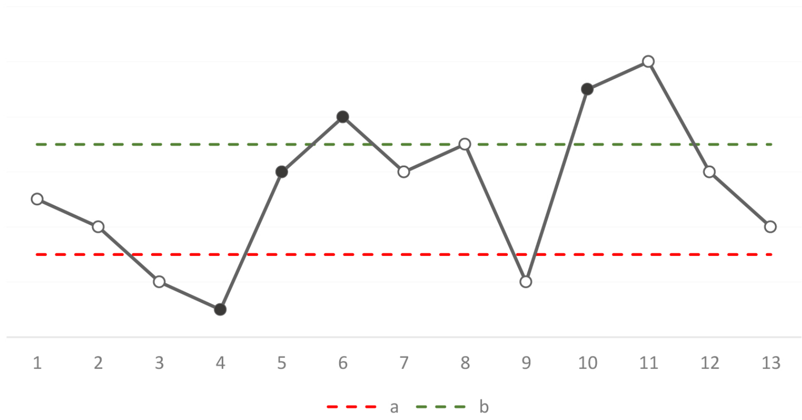


Abbildung 7.1: Aufkreuzungen

Lemma 7.20 (Doobs[2] Aufkreuzungslemma)**.** *Sei X ein Supermartingal. Dann gilt für alle $n \in \mathbb{N}$*

$$\mathbb{E}[U_n^{(a,b)}] \leq \frac{1}{b-a}\mathbb{E}[(X_n - a)^-].$$

Beweis. Nach Lemma 7.10 ist $Y = C \bullet X$ ein Supermartingal. Also gilt

$$0 = \mathbb{E}[Y_0] \geq \mathbb{E}[Y_n] \geq (b-a)\mathbb{E}\big[U_n^{(a,b)}\big] - \mathbb{E}[(X_n - a)^-].$$

□

Satz 7.21 (Doobs (Super-) Martingalkonvergenzsatz)**.** *Ist $(X_n)_n$ ein Supermartingal mit $\sup_n \mathbb{E}[X_n^-] < \infty$, dann gibt es ein $X_\infty \in \mathcal{L}^1(\mathbb{P})$ mit $X_n \to X_\infty$ f.s.*

Beweis. Sei $a < b$. Es gilt $U_n^{(a,b)} \nearrow U_\infty^{(a,b)} = \sup_m U_m^{(a,b)}$ nach Konstruktion. Da

$$\mathbb{E}\big[U_\infty^{(a,b)}\big] = \lim_{n\to\infty} \mathbb{E}\big[U_n^{(a,b)}\big] \leq \sup_n \frac{1}{b-a}\mathbb{E}[(X_n - a)^-] \leq \sup_n \frac{1}{b-a}(\mathbb{E}[X_n^-] + |a|) < \infty$$

ist $U_\infty^{(a,b)} < \infty$ f.s. Für

$$O_{a,b} := \left\{\liminf_{n\to\infty} X_n < a\right\} \cap \left\{\limsup_{n\to\infty} X_n > b\right\} \subset \left\{U_\infty^{(a,b)} = \infty\right\}$$

gilt also $\mathbb{P}(O_{a,b}) = 0$. Damit folgt

$$\mathbb{P}\Big(\liminf_{n\to\infty} X_n < \limsup_{n\to\infty} X_n\Big) = \mathbb{P}\left(\bigcup_{\substack{a<b\\a,b\in\mathbb{Q}}} O_{a,b}\right) = 0.$$

Mit $X_\infty := \limsup_n X_n$ gilt also $X_n \to X_\infty$ f.s. Es bleibt $X_\infty \in \mathcal{L}^1(\mathbb{P})$ zu zeigen. Es gilt:

$$\mathbb{E}[X_\infty^-] = \mathbb{E}[\liminf_{n\to\infty} X_n^-] \overset{\text{Fatou}}{\leq} \liminf_{n\to\infty} \mathbb{E}[X_n^-] < \infty$$

[2] Joseph L. Doob, 1910-2004

nach Voraussetzung und

$$\mathbb{E}[X_\infty^+] = \mathbb{E}[\liminf_{n\to\infty} X_n^+] \le \liminf_{n\to\infty} \mathbb{E}[X_n^+] = \liminf_{n\to\infty}(\mathbb{E}[X_n^-]+\mathbb{E}[X_n]) \le \mathbb{E}[X_0]+\sup_n \mathbb{E}[X_n^-] < \infty.$$

□

Bemerkung 7.22. 1. Die analoge Aussage von Satz 7.21 gilt für ein Submartingal $(X_n)_n$ mit $\sup_n \mathbb{E}[X_n^+] < \infty$.

2. In der Situation von Satz 7.21 liegt im Allgemeinen keine $\mathcal{L}^1$-Konvergenz vor, insbesondere ist $\mathbb{E}[X_\infty] \neq \lim_{n\to\infty} \mathbb{E}[X_n]$ möglich, betrachte zum Beispiel die symmetrische gewöhnliche Irrfahrt startend in 1, gestoppt bei Erreichen der 0.

7.3 Gleichgradig integrierbare Martingale und optionales Stoppen

Satz 7.23. *Sei $(X_n)_n$ ein gleichgradig integrierbares Supermartingal. Dann existiert $X_\infty \in \mathcal{L}^1(\mathbb{P})$ mit $X_n \to X_\infty$ f.s. und in $\mathcal{L}^1(\mathbb{P})$. Es gilt*

$$\mathbb{E}[X_\infty \mid \mathscr{F}_n] \le X_n \text{ f.s. für alle } n \in \mathbb{N}.$$

(Analog gilt $\mathbb{E}[X_\infty \mid \mathscr{F}_n] \ge X_n$, falls $(X_n)_n$ ein gleichgradig integrierbares Submartingal, und $\mathbb{E}[X_\infty \mid \mathscr{F}_n] = X_n$, falls $(X_n)_n$ ein gleichgradig integrierbares Martingal.)

Beweis. Die Existenz von X_∞ mit $X_n \to X_\infty$ f.s. folgt aus Satz 7.21. Aufgrund der gleichgradigen Integrierbarkeit gilt $\mathbb{E}[|X_n - X_\infty|] \xrightarrow[n\to\infty]{} 0$. Weiter gilt für $n \ge m$:

$$\begin{aligned}\mathbb{E}\big[|\mathbb{E}[X_\infty \mid \mathscr{F}_m] - \mathbb{E}[X_n \mid \mathscr{F}_m]|\big] = \mathbb{E}\big[|\mathbb{E}[X_\infty - X_n \mid \mathscr{F}_m]|\big] &\le \mathbb{E}\big[\mathbb{E}[|X_\infty - X_n| \mid \mathscr{F}_m]\big] \\ &= \mathbb{E}[|X_\infty - X_n|] \xrightarrow[n\to\infty]{} 0\end{aligned}$$

und damit ist

$$\mathbb{E}\big[(\mathbb{E}[X_\infty \mid \mathscr{F}_m] - X_m)^+\big] \le \underbrace{\mathbb{E}\big[(\mathbb{E}[X_\infty \mid \mathscr{F}_m] - \mathbb{E}[X_n \mid \mathscr{F}_m])^+\big]}_{\to 0} + \mathbb{E}\big[(\underbrace{\mathbb{E}[X_n \mid \mathscr{F}_m]}_{\le X_m \text{ für } n>m} - X_m)^+\big].$$

Also gilt $\mathbb{E}[X_\infty \mid \mathscr{F}_m] \le X_m$ f.s. □

Satz 7.24. *Sei $Y \in \mathcal{L}^1(\mathbb{P})$ und $(\mathscr{F}_n)_n$ eine Filtration. Sei $\mathscr{F}_\infty := \sigma(\mathscr{F}_n, n \in \mathbb{N})$ und $X_n := \mathbb{E}[Y \mid \mathscr{F}_n]$. Dann ist $(X_n)_n$ ein gleichgradig integrierbares Martingal und es gilt*

$$X_n \xrightarrow[n\to\infty]{} X_\infty := \mathbb{E}[Y \mid \mathscr{F}_\infty] \quad \text{f.s. und in } \mathcal{L}^1(\mathbb{P}).$$

Lemma 7.25. *$Y \in \mathcal{L}^1(\mathbb{P})$ und $\mathscr{F}_n \subset \mathscr{F}$ σ-Algebren, dann ist $(\mathbb{E}[Y \mid \mathscr{F}_n])_{n\in\mathbb{N}}$ gleichgradig integrierbar.*

Beweis. Es existiert (mit Bem. 3.30 und Satz 3.33) ein $h\colon [0,\infty) \to [0,\infty)$ konvex und monoton wachsend mit $\frac{h(x)}{x} \xrightarrow[x\to\infty]{} \infty$ und $\mathbb{E}[h(|Y|)] < \infty$.

Es gilt mit der (bedingten) Jensen-Ungleichung (Satz 6.7)

$$\sup_n \mathbb{E}\big[h(|\mathbb{E}[Y \mid \mathscr{F}_n]|)\big] \le \sup_n \mathbb{E}\big[\mathbb{E}[h(|Y|) \mid \mathscr{F}_n]\big] = \mathbb{E}[h(|Y|)] < \infty.$$

$\square$

Beweis von Satz 7.24. $(X_n)_n$ ist ein Martingal und nach Lemma 7.25 gleichgradig integrierbar. Nach Satz 7.23 gilt

$$X_n \xrightarrow[n\to\infty]{} X_\infty := \limsup_{m\to\infty} X_m \quad \text{f.s. und in } \mathcal{L}^1(\mathbb{P}),$$

X_∞ ist $\mathscr{F}_\infty$-messbar.

Es bleibt $X_\infty = \mathbb{E}[Y \mid \mathscr{F}_\infty]$ zu zeigen. Ohne Einschränkung sei $Y \ge 0$ (ansonsten betrachte Y^+, Y^- separat). Insbesondere ist dann $X_\infty \ge 0$ f.s.

Für $A \in \mathscr{F}_\infty$ sind $\mu_1(A) := \mathbb{E}[X_\infty \mathbf{1}_A]$ und $\mu_2(A) := \mathbb{E}[Y\mathbf{1}_A]$ endliche Maße auf $(\Omega, \mathscr{F}_\infty, \mathbb{P})$. Sei $A \in \mathscr{F}_m \subset \mathscr{F}_\infty$. Dann gilt

$$\mu_1(A) = \mathbb{E}[X_\infty \mathbf{1}_A] = \lim_{n\to\infty} \mathbb{E}[X_n \mathbf{1}_A] \overset{n\ge m}{=} \mathbb{E}[Y\mathbf{1}_A] = \mu_2(A).$$

Da $\bigcup_{m\in\mathbb{N}} \mathscr{F}_m$ ein schnittstabiler Erzeuger von $\mathscr{F}_\infty$ ist, folgt $\mu_1 = \mu_2$ mit Satz 1.18. $\square$

Bemerkung 7.26. Für ein gleichgradig integrierbares Martingal $(X_n)_n$ gilt

$$\mathbb{E}[X_\infty \mid \mathscr{F}_n] = X_n \text{ für alle } n \in \mathbb{N}.$$

Solche Martingale heißen *Doob'sche Martingale.*

Lemma 7.27. *Sei $(X_n)_n$ ein Supermartingal und T eine Stoppzeit mit $T \le m$ f.s. für ein $m \in \mathbb{N}$. Dann ist $X_T \in \mathcal{L}^1(\mathbb{P})$ und es gilt $\mathbb{E}[X_m \mid \mathscr{F}_T] \le X_T$ f.s. Im Falle eines Martingals gilt Gleichheit.*

Beweis. Nach Lemma 7.18 ist $(X_{T\wedge n})_n$ ein Supermartingal. Es gilt $X_T \in \mathcal{L}^1(\mathbb{P})$, denn $X_T = X_{T\wedge m}$ f.s. Für $A \in \mathscr{F}_T$ gilt

$$\mathbb{E}[X_m \mathbf{1}_A] = \sum_{n=0}^m \mathbb{E}\big[X_m \mathbf{1}_{\underbrace{A\cap\{T=n\}}_{\in\mathscr{F}_n}}\big] \le \sum_{n=0}^m \mathbb{E}\big[X_n \mathbf{1}_{A\cap\{T=n\}}\big] = \mathbb{E}\left[\left(\sum_{n=0}^m X_n \mathbf{1}_{\{T=n\}}\right)\mathbf{1}_A\right] = \mathbb{E}[X_T \mathbf{1}_A].$$

$\square$

Lemma 7.28. *Ist $(X_n)_n$ ein gleichgradig integrierbares Martingal, so ist $\{X_T \mid T \text{ Stoppzeit}\}$ gleichgradig integrierbar.*

Beweis. Da $(X_n)_n$ gleichgradig integrierbar ist, existiert (mit Satz 3.33) ein $h\colon[0,\infty)\to[0,\infty)$ konvex und monoton wachsend mit $\frac{h(x)}{x}\xrightarrow[x\to\infty]{}\infty$ und $\mathbb{E}[h(|X_n|)] =: M < \infty$. Sei T eine Stoppzeit und $n\in\mathbb{N}$. Es gilt

$$\begin{aligned}\mathbb{E}\big[h(|X_T|)\mathbf{1}_{\{T\le n\}}\big] &= \mathbb{E}\big[h(|X_{T\wedge n}|)\mathbf{1}_{\{T\le n\}}\big] \overset{7.27}{=} \mathbb{E}\big[h(|\mathbb{E}[X_n \mid \mathscr{F}_{T\wedge n}]|)\mathbf{1}_{\{T\le n\}}\big]\\ &\le \mathbb{E}\big[h(\mathbb{E}[|X_n| \mid \mathscr{F}_{T\wedge n}])\mathbf{1}_{\{T\le n\}}\big] = \mathbb{E}\big[\mathbb{E}[h(|X_n|)\mathbf{1}_{\{T\le n\}} \mid \mathscr{F}_{T\wedge n}]\big]\\ &\le M.\end{aligned}$$

Mit $n\to\infty$ folgt $\mathbb{E}\big[h(|X_T|)\mathbf{1}_{\{T<\infty\}}\big]\le M$, das heißt

$$\sup_{T\ \text{Stoppzeit}} \mathbb{E}[h(|X_T|)] \le 2M < \infty.$$

□

Satz 7.29 (optional-sampling-Theorem)**.** *Es sei $(X_n)_n$ ein gleichgradig integrierbares Martingal, $X_\infty := \lim X_n$ und T eine Stoppzeit. Dann ist $X_T\in\mathcal{L}^1(\mathbb{P})$ und es gilt $\mathbb{E}[X_\infty \mid \mathscr{F}_T] = X_T$ f.s. Insbesondere gilt $\mathbb{E}[X_T]=\mathbb{E}[X_\infty]=\mathbb{E}[X_0]$. Ist S eine Stoppzeit mit $S\le T$, so gilt $\mathbb{E}[X_T \mid \mathscr{F}_S] = X_S$ f.s.*

Beweis. Nach Satz 7.24 gilt $\mathbb{E}[X_\infty \mid \mathscr{F}_m] = X_m$ f.s. und nach Lemma 7.27 ist

$$\mathbb{E}[X_m \mid \mathscr{F}_{T\wedge m}] = X_{T\wedge m} \ \text{ f.s.}$$

Also ist

$$\mathbb{E}[X_\infty \mid \mathscr{F}_{T\wedge m}] = \mathbb{E}[\mathbb{E}[X_\infty \mid \mathscr{F}_m] \mid \mathscr{F}_{T\wedge m}] = X_{T\wedge m} \ \text{ f.s.}$$

Sei nun $A\in\mathscr{F}_T$. Dann ist $A\cap\{T\le m\}\in\mathscr{F}_{T\wedge m}$, denn für $n\in\mathbb{N}_0$ gilt $A\cap\{T\le m\}\cap\{T\le n\} = A\cap\{T\le m\wedge n\}\in\mathscr{F}_{m\wedge n}\subset\mathscr{F}_n$. Somit gilt

$$\begin{aligned}\mathbb{E}\big[X_\infty\mathbf{1}_{A\cap\{T\le m\}}\big] &= \mathbb{E}[\mathbb{E}[X_\infty\mathbf{1}_{A\cap\{T\le m\}} \mid \mathscr{F}_{T\wedge m}]]\\ &= \mathbb{E}\big[X_{T\wedge m}\mathbf{1}_{A\cap\{T\le m\}}\big] = \mathbb{E}\big[X_T\mathbf{1}_{A\cap\{T\le m\}}\big].\end{aligned} \tag{7.2}$$

Sei ohne Einschränkung $X_\infty\ge 0$ (sonst betrachte X_∞^+, X_∞^- separat). $m\to\infty$ in (7.2) mit monotoner Konvergenz liefert

$$\mathbb{E}\big[X_\infty\mathbf{1}_{A\cap\{T<\infty\}}\big] = \mathbb{E}\big[X_T\mathbf{1}_{A\cap\{T<\infty\}}\big].$$

(Alternativ: Verwende dominierte Konvergenz, das erspart die Zerlegung in Positiv- und Negativteil.)

Nach Definition gilt aber auch

$$\mathbb{E}\big[X_\infty\mathbf{1}_{A\cap\{T=\infty\}}\big] = \mathbb{E}\big[X_T\mathbf{1}_{A\cap\{T=\infty\}}\big],$$

d.h. $\mathbb{E}[X_\infty\mathbf{1}_A] = \mathbb{E}[X_T\mathbf{1}_A]$.

Sei $S\le T$ eine Stoppzeit. Dann ist $\mathscr{F}_S\subset\mathscr{F}_T$, also gilt

$$\mathbb{E}[X_T \mid \mathscr{F}_S] = \mathbb{E}\big[\mathbb{E}[X_\infty \mid \mathscr{F}_T] \mid \mathscr{F}_S\big] = \mathbb{E}[X_\infty \mid \mathscr{F}_S] = X_S \quad \text{f.s.}$$

□

Bemerkung und Definition 7.30. Sei $(X_n)_n$ ein adaptierter Prozess mit $X_n \in \mathcal{L}^1(\mathbb{P})$ für $n \in \mathbb{N}_0$. Dann ist $X_n = M_n + A_n$ mit

$$M_0 := X_0, \quad M_n := X_0 + \sum_{k=1}^{n} \big(X_k - \mathbb{E}[X_k \mid \mathscr{F}_{k-1}] \big),$$

$$A_0 := 0, \quad A_n := \sum_{k=1}^{n} \big(\mathbb{E}[X_k \mid \mathscr{F}_{k-1}] - X_{k-1} \big).$$

wobei $(M_n)_n$ ein Martingal und $(A_n)_n$ previsibel ist. Die Darstellung $X = M + A$ als Summe eines Martingals M und eines previsiblen Prozesses A mit $A_0 = 0$ heißt *Doob-Zerlegung*, sie ist f.s. eindeutig.

$(X_n)_n$ ist genau dann ein Super- bzw. Submartingal, wenn $(A_n)_n$ nicht-wachsend bzw. nicht-fallend ist.

Beweis. Wir zeigen nur die Eindeutigkeit. Angenommen $X = M + A = M' + A'$. Sei

$$\tilde{M}_n := M_n - M'_n = A'_n - A_n.$$

Dann ist $(\tilde{M}_n)_n$ ein previsibles Martingal mit $\tilde{M}_0 = 0$. Also ist $\tilde{M}_n \equiv \tilde{M}_0 \equiv 0$ f.s., denn

$$\tilde{M}_{n-1} = \mathbb{E}[\tilde{M}_n \mid \mathscr{F}_{n-1}] = \tilde{M}_n \text{ f.s.}$$

□

Satz 7.31. *Sei $(X_n)_n$ ein gleichgradig integrierbares Supermartingal und seien S, T Stoppzeiten mit $S \le T$. Dann gilt $\mathbb{E}[X_T \mid \mathscr{F}_S] \le X_S$ f.s.*

Beweis. Sei $X_n = M_n + A_n$ die Doob-Zerlegung. Dann gilt $A_n \searrow A_\infty \le 0$. Es ist

$$\mathbb{E}[|A_n|] = \mathbb{E}[-A_n] = \mathbb{E}[M_n - X_n] = \mathbb{E}[M_n - M_0 + X_0 - X_n] \le \mathbb{E}[|X_n| + \mathbb{E}[|X_0|]] \le C$$

für alle n mit einer geeigneten Konstante $C < \infty$.

Somit ist $(A_n)_n$ gleichgradig integrierbar und damit auch $(M_n)_n = (X_n - A_n)_n$. Also gilt

$$\mathbb{E}[X_T \mid \mathscr{F}_S] = \underbrace{\mathbb{E}[M_T \mid \mathscr{F}_S]}_{=M_S \text{ f.s.}} + \mathbb{E}[A_T \mid \mathscr{F}_S] \le M_S + \mathbb{E}[A_S \mid \mathscr{F}_S] = M_S + A_S = X_S.$$

□

7.4 $\mathcal{L}^2$-Martingale

Bemerkung 7.32. Sei $(X_n)_n$ ein Martingal und $\varphi: \mathbb{R} \to \mathbb{R}$ konvex, sodass $\mathbb{E}[\varphi(X_n)]$ für alle n existiert. Dann ist $(\varphi(X_n))_n$ ein Submartingal, denn

$$\mathbb{E}[\varphi(X_{n+1}) \mid \mathscr{F}_n] \ge \varphi(\mathbb{E}[X_{n+1} \mid \mathscr{F}_n]) = \varphi(X_n).$$

Die Aussage gilt ebenso, wenn $(X_n)_n$ ein Submartingal und φ konvex und nicht fallend ist.

Bericht 7.33. Eine Funktion $f:\mathbb{R}\to\mathbb{R}$ heißt *harmonisch*, wenn sie die Mittelwerteigenschaft besitzt:

$$f(x) = \frac{1}{2a}\int_{x-a}^{x+a} f(y)\,dy \quad \forall x\in\mathbb{R}, a>0$$

(man fordert auch, dass f lokal beschränkt und messbar ist, so dass die Integrale stets existieren).

$f:\mathbb{R}\to\mathbb{R}$ heißt *subharmonisch*, wenn $f(x)\le\frac{1}{2a}\int_{x-a}^{x+a} f(y)\,dy$ und *superharmonisch*, wenn $f(x)\ge\frac{1}{2a}\int_{x-a}^{x+a} f(y)\,dy$ für alle $x\in\mathbb{R}, a>0$.

Man kann zeigen: Auf $\mathbb{R}$ sind genau die konvexen Funktionen subharmonisch, zusammen mit Bemerkung 7.32 motiviert dies den Namen Submartingal (und entsprechend auch den Namen Supermartingal).

Beobachtung 7.34. Sei $(X_n)_n$ ein quadratintegrables Martingal, d.h. $X_n\in\mathcal{L}^2(\mathbb{P})$ für alle $n\in\mathbb{N}$. Dann gilt $\mathbb{E}[(X_k-X_l)(X_l-X_m)]=0$ für alle $0\le m\le l\le k$, denn

$$\begin{aligned}\mathbb{E}[(X_k-X_l)(X_l-X_m)] &= \mathbb{E}\big[\mathbb{E}[(X_k-X_l)(X_l-X_m)\mid\mathscr{F}_l]\big]\\ &= \mathbb{E}\big[(X_l-X_m)(\mathbb{E}[X_k\mid\mathscr{F}_l]-X_l)\big]\\ &= \mathbb{E}[(X_l-X_m)(X_l-X_l)] = \mathbb{E}[0]=0.\end{aligned}$$

Erinnerung (vgl. Beob. 3.27)**.** $\|X\|_2 := \sqrt{\mathbb{E}[X^2]}$ ist eine Norm und $\langle X,Y\rangle := \mathbb{E}[XY]$ ist ein Skalarprodukt, $\mathcal{L}^2(\mathbb{P})$ ist ein Hilbertraum (wenn man $\mathbb{P}$-f.s.-Gleichheit „herausfaktorisiert“).

Man spricht Beob. 7.34 auch aus als: „Martingalinkremente über disjunkte Zeitintervalle sind orthogonal.“

Lemma und Definition 7.35. *Sei $(X_n)_n$ ein quadratintegrables Martingal.*

$$A_n := \sum_{k=1}^{n}\mathbb{E}[(X_k-X_{k-1})^2\mid\mathscr{F}_{k-1}]$$

ist der eindeutig bestimmte previsible Prozess mit $A_0:=0$, sodass $(X_n^2-A_n)_n$ ein Martingal ist. Man schreibt auch $(\langle X\rangle_n)_n=(A_n)_n$. $\langle X\rangle$ heißt quadratische Variation *von X. (In der Literatur werden auch folgende Namen verwendet: Wachsender Prozess, previsible quadratische Variation, Spitzklammerprozess von X.)*

Beweis. Es gilt:

$$\begin{aligned}\mathbb{E}[X_{n+1}^2-A_{n+1}\mid\mathscr{F}_n] &= \mathbb{E}[(X_{n+1}-X_n)^2+2(X_{n+1}-X_n)X_n+X_n^2-A_{n+1}\mid\mathscr{F}_n]\\ &= \underbrace{\mathbb{E}[(X_{n+1}-X_n)^2\mid\mathscr{F}_n]-A_{n+1}}_{=-A_n}+X_n^2+2X_n\underbrace{\mathbb{E}[X_{n+1}-X_n\mid\mathscr{F}_n]}_{=0}\\ &= X_n^2-A_n.\end{aligned}$$

Die Eindeutigkeit folgt aus der Eindeutigkeit der Doob-Zerlegung (Bem. und Def. 7.30). □

Insbesondere gilt also:

$$\mathbb{E}[X_n^2] = \mathbb{E}[X_0^2] + \mathbb{E}[\langle X\rangle_n] = \mathbb{E}[X_0^2] + \sum_{k=1}^{n} \mathbb{E}[(X_k - X_{k-1})^2]$$

und

$$\sup_n \mathbb{E}[X_n^2] < \infty \qquad \Longleftrightarrow \qquad \sup_n \mathbb{E}[\langle X\rangle_n] < \infty.$$

Satz 7.36. *Sei $(M_n)_n$ ein $\mathcal{L}^2$-Martingal. Dann gilt:*

i) $\{\langle M\rangle_\infty < \infty\} \overset{f.s.}{\subset} \left\{\lim_{n\to\infty} M_n \text{ existiert}\right\}$.

ii) Wenn $|M_n - M_{n-1}| \le c$ für alle n für ein $c < \infty$, so gilt auch

$$\left\{\lim_{n\to\infty} M_n \text{ existiert}\right\} \overset{f.s.}{\subset} \{\langle M\rangle_\infty < \infty\}.$$

iii) $\{\langle M\rangle_\infty = \infty\} \overset{f.s.}{\subset} \left\{\frac{M_n}{\langle M\rangle_n} \xrightarrow[n\to\infty]{} 0\right\}$.

Bemerkung 7.37. iii) impliziert das Starke Gesetz der großen Zahlen für $M_n = Y_1 + \ldots + Y_n$ mit Y_i unabhängig und identisch verteilt mit $\mathbb{E}[Y_1] = 0$ und $\mathrm{Var}[Y_1] < \infty$.

Lemma 7.38 (Kroneckers Lemma)**.** *Sei $(x_n)_n \subset \mathbb{R}$ mit $s_k = \sum_{n=1}^k x_n \to s_\infty \in \mathbb{R}$. Ist $0 \le b_n \nearrow \infty$, dann gilt $\frac{1}{b_n}\sum_{k=1}^n b_k x_k \xrightarrow[n\to\infty]{} 0$.*

Beweis von Satz 7.36. Sei $k \in \mathbb{R}^+$.

$$S_k := \inf\{n \in \mathbb{N}_0 \mid \{M\}_{n+1} > k\}$$

ist eine Stoppzeit. Nach Lemma 7.18 und Lemma 7.35 ist $(M_{n\wedge S_k}^2 - \langle M\rangle_{n\wedge S_k})_n$ ein Martingal, also gilt

$$\sup_n \mathbb{E}[M_{n\wedge S_k}^2] = \mathbb{E}[M_0] + \underbrace{\sup_n \mathbb{E}[\langle M\rangle_{n\wedge S_k}]}_{\le k} \le \infty,$$

das heißt $(M_{n\wedge S_k})_n$ ist $\mathcal{L}^2$-beschränkt. Also existiert $\lim_n M_{n\wedge S_k}$ f.s. für jedes $k \in \mathbb{R}^+$. Es gilt $\{\langle M\rangle_\infty < \infty\} = \bigcup_{k\in\mathbb{N}} \{S_k = \infty\}$ und $\bigcup_{k\in\mathbb{N}} \{S_k = \infty, \lim_n M_{n\wedge S_k} \text{ existiert}\} \subset \{\lim_n M_n \text{ existiert}\}$, damit gilt i).

Sei $K > 0$. $T_K := \inf\{n \in \mathbb{N}_0 \mid |M_n| > K\}$ ist eine Stoppzeit und es gilt

$$\mathbb{E}[\underbrace{M_{n\wedge T_K}^2}_{\le (K+c)^2} - \langle M\rangle_{n\wedge T_K}] = \mathbb{E}[M_0^2].$$

Also folgt mit monotoner Konvergenz

$$\mathbb{E}[\langle M\rangle_{T_K}] = \sup_n \mathbb{E}[\langle M\rangle_{n\wedge T_K}] \le \mathbb{E}[M_0^2] + (K+c)^2 < \infty.$$

Demnach ist $\langle M\rangle_{T_K} < \infty$ f.s. und es gilt

$$\mathbb{P}\left(\underbrace{\left\{\lim_{n\to\infty} M_n \text{ existiert}\right\}}_{\subset\bigcup_{k\in\mathbb{N}}\{T_k=\infty\}} \cap \{\langle M\rangle_\infty = \infty\}\right)$$

$$\leq \mathbb{P}\left(\bigcup_{K\in\mathbb{N}} \{T_K = \infty\} \cap \{\langle M\rangle_\infty = \infty\} \cap \{\langle M\rangle_{T_K} < \infty\}\right) = 0.$$

Also gilt *ii*).

Weiter sei

$$W_n := \sum_{k=1}^n \frac{M_k - M_{k-1}}{1 + \langle M\rangle_k} = ((1 + \langle M\rangle)^{-1} \bullet M)_n.$$

Nach Lemma 7.10 ist $(W_n)_n$ ein Martingal und es gilt

$$\begin{aligned}\mathbb{E}[(W_n - W_{n-1})^2 \mid \mathscr{F}_{n-1}] &= \frac{1}{(1 + \langle M\rangle_n)^2}\mathbb{E}[(M_n - M_{n-1})^2 \mid \mathscr{F}_{n-1}] = \frac{\langle M\rangle_n - \langle M\rangle_{n-1}}{(1 + \langle M\rangle_n)^2}\\ &\leq \frac{(\langle M\rangle_n + 1) - (\langle M\rangle_{n-1} + 1)}{(1 + \langle M\rangle_n)(1 + \langle M\rangle_{n-1})} = \frac{1}{1 + \langle M\rangle_{n-1}} - \frac{1}{1 + \langle M\rangle_n}.\end{aligned}$$

Also gilt

$$\langle W\rangle_\infty = \lim_{N\to\infty} \langle W\rangle_N \leq \limsup_{N\to\infty} \sum_{n=1}^N \left(\frac{1}{1 + \langle M\rangle_{n-1}} - \frac{1}{1 + \langle M\rangle_n}\right) \leq 1$$

und somit $W_n \to W_\infty$ f.s. nach i). Wähle nun $b_n = 1 + \langle M\rangle_n$ und $x_n = \frac{M_n - M_{n-1}}{1+\langle M\rangle_n}$ in Lemma 7.38, dann folgt $\sum_{k=1}^n b_k x_k = \sum_{k=1}^n M_k - M_{k-1} = M_n - M_0$, d.h. *iii*) gilt. □

Im Folgenden sei $(X_n)_n$ ein stochastischer Prozess mit Werten in $\mathbb{R}$ und

$$X_n^* := \max_{0\leq k\leq n} X_k, \quad |X|_n^* := \max_{0\leq k\leq n} |X_k|.$$

Lemma 7.39. *Sei $(X_n)_n$ ein Submartingal und $\lambda \geq 0$. Dann gilt*

$$\lambda\mathbb{P}(X_n^* \geq \lambda) \leq \mathbb{E}\big[X_n \mathbf{1}_{\{X_n^*\geq\lambda\}}\big] \leq \mathbb{E}\big[|X_n| \mathbf{1}_{\{X_n^*\geq\lambda\}}\big] \ (\leq \mathbb{E}[|X_n|]).$$

Beweis. Sei $T := \inf\{k \in \mathbb{N}_0 \mid X_k \geq \lambda\} \wedge n$. T ist eine beschränkte Stoppzeit und es gilt mit Korollar 7.19

$$\begin{aligned}\mathbb{E}\big[X_n \mathbf{1}_{\{X_n^*\geq\lambda\}}\big] + \mathbb{E}\big[X_n \mathbf{1}_{\{X_n^*<\lambda\}}\big] = \mathbb{E}[X_n] \overset{7.19}{\geq} \mathbb{E}[X_T] &= \mathbb{E}\big[X_T \mathbf{1}_{\{X_n^*\geq\lambda\}}\big] + \mathbb{E}\big[X_T \mathbf{1}_{\{X_n^*<\lambda\}}\big]\\ &\geq \lambda\mathbb{P}(X_n^* \geq \lambda) + \mathbb{E}\big[X_n \mathbf{1}_{\{X_n^*<\lambda\}}\big].\end{aligned}$$

□

Satz 7.40 (Doobs $\mathcal{L}^p$-Ungleichungen)**.** *Sei $(X_n)_n$ ein Martingal oder ein nichtnegatives Submartingal.*

i) Für $p \geq 1$ *und* $\lambda > 0$ *gilt* $\lambda^p \mathbb{P}(|X|_n^* \geq \lambda) \leq \mathbb{E}[|X_n|^p]$.

ii) Für $p > 1$ *gilt* $\mathbb{E}[|X_n|^p] \leq \mathbb{E}[(|X|_n^*)^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_n|^p]$.

Bemerkung. Für ein $\mathcal{L}^2$-Martingal gestattet dies, $\mathbb{E}[(|X|_n^*)^2]$ durch $\mathbb{E}[\langle X\rangle_n]$ zu kontrollieren, denn $\mathbb{E}[X_n^2] = \mathbb{E}[X_0^2] + \mathbb{E}[\langle X\rangle_n]$.

Beweis von Satz 7.40. i) $(|X_n|^p)_n$ ist ein Submartingal, also folgt die Aussage durch Anwenden von Lemma 7.39 auf $(|X_n|^p)_n$.

ii) Sei $c > 0$. Es gilt:

$$\begin{aligned}
\mathbb{E}[(|X|_n^* \wedge c)^p] &= \mathbb{E}\left[\int_0^{|X|_n^* \wedge c} p\lambda^{p-1}\,d\lambda\right] = \mathbb{E}\left[\int_0^c p\lambda^{p-1}\mathbf{1}_{\{|X|_n^* \geq \lambda\}}\,d\lambda\right] \\
&\overset{\text{Fubini}}{=} \int_0^c p\lambda^{p-1}\mathbb{P}(|X|_n^* \geq \lambda)\,d\lambda \leq \int_0^c p\lambda^{p-1}\frac{1}{\lambda}\mathbb{E}[|X_n|\mathbf{1}_{\{|X|_n^* \geq \lambda\}}]\,d\lambda \\
&\overset{\text{Fubini}}{=} p\mathbb{E}\left[|X_n|\int_0^{c\wedge |X|_n^*} \lambda^{p-2}\,d\lambda\right] = \frac{p}{p-1}\mathbb{E}[|X_n|(|X|_n^* \wedge c)^{p-1}] \\
&\overset{\text{Hölder}}{\leq} \frac{p}{p-1}E[(|X|_n^* \wedge c)^p]^{\frac{p-1}{p}}\mathbb{E}[|X_n|^p]^{\frac{1}{p}}.
\end{aligned}$$

Für $c \to \infty$ folgt mit monotoner Konvergenz

$$\mathbb{E}[(|X|_n^*)^p] \leq \frac{p}{p-1}\mathbb{E}[(|X|_n^*)^p]^{\frac{p-1}{p}}\mathbb{E}[|X_n|^p]^{\frac{1}{p}}$$

und durch Umstellen und Potenzieren dieser Ungleichung erhält man

$$\mathbb{E}[(|X|_n^*)^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_n|^p].$$

□

Korollar 7.41 (eine Form der Kolmogorov-Ungleichung)**.** $X_1, X_2, \dots \in \mathcal{L}^2(\mathbb{P})$ *unabhängig,* $\mathbb{E}[X_n] = 0$*,* $S_n := X_1 + \dots + X_n$ *(*$S_0 := 0$*). Es gilt für* $x > 0$

$$\mathbb{P}\Big(\max\{|S_k| : 1 \leq k \leq n\} \geq x\Big) \leq \frac{\mathrm{Var}[S_n]}{x^2}.$$

Beweis. $(S_n)_n$ ist ein Martingal (siehe Bsp. 7.7, i)), Satz 7.40, *i)* liefert

$$\mathbb{P}\Big(\max\{|S_k| : 1 \leq k \leq n\} \geq x\Big) \leq \frac{\mathbb{E}[S_n^2]}{x^2}.$$

□

Korollar 7.42. $X_1, X_2, \dots$ *u.a. mit* $\mathbb{E}[X_n] = 0$*,* $c := \sup_n \mathrm{Var}[X_n] < \infty$*,* $S_n = X_1 + \dots + X_n$*, dann gilt für* $\varepsilon > 0$

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{n}(\log n)^{1/2+\varepsilon}} = 0 \quad \text{f.s.}$$

Beweis. Sei $\ell(n) := \sqrt{n}\,(\log n)^{1/2+\varepsilon}$, $k_n := 2^n$, für $\delta > 0$ sei

$$A_{n,\delta} := \Big\{ \max_{k \le k_n} |S_k| > \delta \ell(k_n) \Big\}.$$

Nach Kor. 7.41 ist

$$\sum_{n=1}^{\infty} \mathbb{P}(A_{n,\delta}) \le \sum_{n=1}^{\infty} \frac{1}{\delta^2 \ell(k_n)^2} \mathrm{Var}[S_{k_n}] \le \sum_{n=1}^{\infty} \frac{c k_n}{\delta^2 \ell(k_n)^2} = \frac{c}{\delta^2} \sum_{n=1}^{\infty} \frac{1}{\big(\log(2^n)\big)^{1+2\varepsilon}} < \infty,$$

mit Borel-Cantelli also

$$\limsup_{n\to\infty} \frac{1}{\ell(k_n)} \max\Big\{ |S_k| : k \le k_n \Big\} \le \delta \quad \text{f.s.}$$

Mit $\delta \downarrow 0$ folgt $\max\Big\{|S_k| : k \le k_n\Big\}/\ell(k_n) \to 0$ f.s. für $n \to \infty$ und daher auch

$$\limsup_{m\to\infty} \frac{|S_m|}{\ell(m)} \le \limsup_{m\to\infty} \underbrace{\frac{\ell(k_{\lceil \log_2 m \rceil})}{\ell(m)}}_{\text{ist beschr.}} \frac{\max\Big\{|S_k| : k \le k_{\lceil \log_2 m\rceil}\Big\}}{\ell(k_{\lceil \log_2 m \rceil})} = 0 \quad \text{f.s.}$$

□

Bericht 7.43. Seien $X_1, X_2, \dots$ u.i.v. mit $\mathbb{E}[X_1] = 0$, $\sigma^2 := \mathrm{Var}[X_1] \in (0,\infty)$, $S_n = X_1 + \dots + X_n$. Für sehr großes n ist

$$\frac{S_n}{\sigma\sqrt{n}} \overset{d}{\approx} \mathcal{N}_{0,1}$$

gemäß dem zentralen Grenzwertsatz (siehe auch das folgende Kapitel 8), d.h. für festes (und großes) n sind die typischen Fluktuationen von S_n von der Ordnung $O(\sqrt{n})$.

Korollar 7.42 zeigt, dass die Fluktuationen von S_n „global" (d.h., wenn wir „irgendwo" auf dem Pfad nachschauen dürfen), höchstens um einen Faktor $(\log n)^{1/2+\varepsilon}$ größer sind als $\sqrt{n}$. Der Faktor aus Kor. 7.42 ist nicht scharf, tatsächlich gilt

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\sigma^2 \log\log n}} = 1 \quad \text{f.s.},$$

das sogenannte Gesetz vom iterierten Logarithmus (siehe z.B. [Kl, Satz 22.9]).

Satz 7.44. *1)* $X_1, X_2, \dots \in \mathcal{L}^2(\mathbb{P})$ *u.a,* $\mathbb{E}[X_n] = 0$, $\sum_{n=1}^{\infty} \mathrm{Var}[X_n] < \infty$, *dann konvergiert* $\sum_{n=1}^{\infty} X_n$ *f.s.*

2) $X_1, X_2, \dots$ *u.a., reellwertige ZVn, es gelte für ein* $a > 0$ *mit* $Y_n := X_n \mathbf{1}_{\{|X_n| \le a\}}$

i) $\sum_{n=1}^{\infty} \mathbb{P}\big(|X_n| > a\big) < \infty$

ii) $\sum_{n=1}^{\infty} \mathbb{E}[Y_n] < \infty$

iii) $\sum_{n=1}^{\infty} \mathrm{Var}[Y_n] < \infty$

Dann konvergiert $\sum_{n=1}^{\infty} X_n$ *f.s.*

Beweis. 1) Sei $S_n := X_1 + \dots + X_n$, für $m < n$ und $\varepsilon > 0$ ist

$$\mathbb{P}\Big(\max_{m \le k \le n} |S_k - S_m| > \varepsilon \Big) \le \varepsilon^{-2} \mathrm{Var}[S_n - S_m] = \varepsilon^{-2} \sum_{k=m}^{n} \mathrm{Var}[X_k]$$

nach Kor. 7.41, mit $n \to \infty$ folgt

$$\mathbb{P}\Big(\max_{k \ge m} |S_k - S_m| > \varepsilon \Big) \le \varepsilon^{-2} \sum_{k=m}^{\infty} \mathrm{Var}[X_k] \quad \big(\underset{m \to \infty}{\longrightarrow} 0 \big).$$

Sei $W_m := \sup_{\ell, k \ge m} |S_k - S_\ell|$, offenbar ist W_m nicht-wachsend in m und

$$\mathbb{P}(W_m \ge 2\varepsilon) \le \mathbb{P}\Big(\max_{k \ge m} |S_k - S_m| > \varepsilon \Big) \underset{m \to \infty}{\longrightarrow} 0$$

somit gilt $W_m \to 0$ f.s. für $m \to \infty$, d.h. $(S_n)_n$ ist f.s. Cauchy-Folge und daher konvergent.

2) 1) zusammen mit *ii)* und *iii)* liefert

$$\sum_{n=1}^{\infty} \big(Y_n - \mathbb{E}[Y_n] \big) \text{ konvergiert f.s} \quad \text{und} \quad \sum_{n=1}^{\infty} \mathbb{E}[Y_n] \text{ konvergiert.}$$

i) und Borel-Cantelli-Lemma liefern $X_n = Y_n$ für $n \ge N_0$ für ein (zufälliges) $N_0 < \infty$, folglich konvergiert auch $\sum_{n=1}^{\infty} X_n$ f.s. □

Beispiel 7.45. Seien $Z_1, Z_2, \dots$ u.i.v. $\sim \mathcal{N}_{0,1}$, $t \in \mathbb{R}$, $X_n := Z_n \frac{1}{n} \sin(n\pi t)$. $\sum_{n=1}^{\infty} X_n$ konvergiert f.s. (mit Satz 7.44, 1)).

7.5 Zum Satz von Radon-Nikodým

Sei $(S, \mathscr{A})$ ein messbarer Raum, μ, ν Maße auf $(S, \mathscr{A})$. ν ist *absolut stetig* bezüglich μ, geschrieben $\nu \ll \mu$, wenn für jedes $A \in \mathscr{A}$ gilt

$$\mu(A) = 0 \quad \Longrightarrow \quad \nu(A) = 0.$$

Wenn ν eine Dichte bezüglich μ besitzt ($\nu = h\mu$, also $\nu(A) = \int \mathbf{1}_A h \, d\mu$ mit einem $h : S \to \overline{\mathbb{R}}_+$ messbar, vgl. Def. 3.16), so gilt offenbar $\nu \ll \mu$.

Satz 7.46 (Eine Form des Satzes von Radon-Nikodým[3])**.** *Sei* $(S, \mathscr{A})$ *separabler messbarer Raum (d.h.* $\mathscr{A} = \sigma(A_1, A_2, \dots)$ *ist abzählbar erzeugt),* μ, ν *endliche Maße auf* $(S, \mathscr{A})$ *mit* $\nu \ll \mu$*. Dann gibt es ein* $h \ge 0$, $h \in \mathcal{L}^1(\mu)$ *mit* $\nu = h\mu$*.*

[3] Johann Radon, 1887–1956; Otton Nikodým, 1887–1974

Beweis. O.E. sei $\mu(S), \nu(S) > 0$, dann sind $P(\cdot) := \mu(\cdot)/\mu(S)$, $Q(\cdot) := \nu(\cdot)/\nu(S)$ W'maße auf $(S, \mathscr{A})$ mit $Q \ll P$ und es genügt zu zeigen, dass $Q = XP$ für ein $X \in \mathcal{L}^1(P)$.

Wir beobachten zunächst:

$$\text{für } \varepsilon > 0 \text{ gibt es } \delta > 0 \text{ mit } \quad P(B) < \delta \ \Rightarrow\ Q(B) < \varepsilon \qquad (*)$$

Wenn $(*)$ nicht gälte, so gäbe es $B_n \in \mathscr{A}$ und ein $\varepsilon_0 > 0$ mit

$$P(B_n) \le 2^{-n},\ n \in \mathbb{N}, \qquad \inf_{n\in\mathbb{N}} Q(B_n) \ge \varepsilon_0.$$

Dann gälte für $C := \limsup_{n\to\infty} B_n = \bigcap\limits_n \bigcup\limits_{m\ge n} B_m$ einerseits $P(C) = 0$ mit Borel-Cantelli, andererseits wegen $\inf_{n\in\mathbb{N}} Q\big(\bigcup_{m\ge n} B_m\big) \ge \varepsilon_0$ auch $Q(C) \ge \varepsilon_0$ im Widerspruch zu $Q \ll P$.

Sei

$$\mathscr{F}_n := \sigma(A_1, \dots, A_n) = \{C_{n,1}, \dots, C_{n,m_n}\}$$

(mit $S = \dot\bigcup_{k=1}^{m_n} C_{n,k}$, $C_{n,k} \ne \varnothing$), für $\omega \in S$ setze

$$X_n(\omega) = \begin{cases} \frac{Q(C_{n,k})}{P(C_{n,k})}, & \omega \in C_{n,k} \text{ und } P(C_{n,k}) > 0, \\ 0, & \omega \in C_{n,k} \text{ und } P(C_{n,k}) = 0. \end{cases}$$

$X_n \ge 0$ ist $\mathscr{F}_n$-messbar, $X_n \in \mathcal{L}^1(P)$ mit $\mathbb{E}_P[X_n] = 1$ und für $B \in \mathscr{F}_n$ (d.h. $B = \dot\bigcup_{j\in J} C_{n,j}$ für ein $J \subset \{1, \dots, m_n\}$) ist

$$\mathbb{E}_P[X_n \mathbf{1}_B] = \sum_{j\in J} \mathbb{E}_P[X_n \mathbf{1}_{C_{n,j}}] = \sum_{j\in J} Q(C_{n,j}) = Q(B),$$

d.h. X_n ist die Dichte von $Q|_{\mathscr{F}_n}$ bzgl. $P|_{\mathscr{F}_n}$.

Offenbar ist $(\mathscr{F}_n)_n$ eine Filtration, für $B' \in \mathscr{F}_n$ ($\subset \mathscr{F}_{n+1}$) ist $\mathbb{E}_P[X_{n+1}\mathbf{1}_{B'}] = Q(B') = \mathbb{E}_P[X_n \mathbf{1}_{B'}]$ nach obigem, d.h.

$$\mathbb{E}_P[X_{n+1} \mid \mathscr{F}_n] = X_n \quad \text{(f.s.)}$$

$(X_n)_n$ ist also ein nicht-negatives (P-)Martingal.

Zeige: $\{X_n, n \in \mathbb{N}\}$ ist gleichgradig (P-)integrierbar.

Zu $\varepsilon > 0$ wähle $\delta > 0$ gemäß $(*)$, setze $M := 1/\delta$, dann gilt für jedes $n \in \mathbb{N}$

$$P(X_n > M) \le \frac{1}{M}\mathbb{E}_P[X_n] = \delta,$$

mit $(*)$ also

$$\mathbb{E}_P\big[X_n \mathbf{1}_{\{X_n \ge M\}}\big] = Q(X_n > M) \le \varepsilon.$$

Mit Satz 7.23 gibt es $X_\infty \in \mathcal{L}^1(P)$ mit $\mathbb{E}_P[X_\infty] = 1$ und $X_n = \mathbb{E}_p[X_\infty \mid \mathscr{F}_n] \to X_\infty$ f.s. und in $\mathcal{L}^1(P)$.

X_∞ leistet das Gewünschte: Das W'maß $\widetilde{Q}(A) := \mathbb{E}_P[X_\infty \mathbf{1}_A]$, $A \in \mathscr{A}$ erfüllt

$$\widetilde{Q}(B) = \mathbb{E}_P[X_\infty \mathbf{1}_B] = \mathbb{E}_P[X_n \mathbf{1}_B] = Q(B) \quad \text{falls } B \in \mathscr{F}_n \text{ für ein } n \in \mathbb{N},$$

d.h. $Q = \widetilde{Q}$ auf $\bigcup_n \mathscr{F}_n$ (und dies ist ein $\cap$-stabiler Erzeuger von $\mathscr{A}$). Mit Satz 1.18 folgt $Q = \widetilde{Q}$. □

Bericht 7.47. 1. Man nennt die Dichte h von $\nu = h\mu$ bezüglich μ auch die Radon-Nikodým-Ableitung (von ν bezüglich μ) und notiert dies auch als $\frac{d\nu}{d\mu} = h$.

2. Satz 7.46 gilt genauso, wenn μ und ν σ-endlich sind (zerlege $S = \dot\bigcup_k S_k$, so dass $\mu(S_k), \nu(S_k) < \infty$, dann gibt es jeweils $\frac{d\mathbf{1}_{S_k}\nu}{d\mathbf{1}_{S_k}\mu}$).

Auch auf die Forderung, dass $\mathscr{A}$ einen abzählbaren Erzeuger besitzt, kann man verzichten, siehe z.B. [Wi, Ch. 14.13].

Kapitel 8

Zum zentralen Grenzwertsatz

Satz 8.1 („Zentraler Grenzwertsatz“). *Seien $X_1, X_2, \dots$ u.i.v. reelle ZVn $\in \mathcal{L}^2$ mit $\sigma^2 = \mathrm{Var}[X_1] \in (0,\infty)$, dann gilt für $-\infty \le a < b \le \infty$*

$$\lim_{n\to\infty} \mathbb{P}\Big(a \le \frac{X_1 + \cdots + X_n - n\mathbb{E}[X_1]}{\sqrt{n\sigma^2}} \le b\Big) = \mathbb{P}(a \le Z \le b) \quad \text{mit } Z \sim \mathcal{N}_{0,1}.$$

Ein „traditioneller“ Zugang zum zentralen Grenzwertsatz verwendet charakteristische Funktionen. Wir diskutieren dies hier etwas skizzenhaft, für Details siehe z.B. [Wi, Ch. 18] oder [Kl, Kap. 15].

Für eine reellwertige ZV X ist die *charakteristische Funktion* definiert als

$$\varphi_X(t) = \mathbb{E}\big[e^{itX}\big], \quad t \in \mathbb{R}$$

wbei $i = \sqrt{-1}$ die imaginäre Einheit ist (wir haben bisher nur reellwertige Funktionen integriert, obiger Erwartungswert ist durch Zerlegung in Real- und Imaginärteil erklärt, d.h. $\mathbb{E}\big[e^{itX}\big] = \mathbb{E}\big[\cos(tX)\big] + i\mathbb{E}\big[\sin(tX)\big]$).

Eine entscheidende Eigenschaft der charakteristischen Funktion ist:

$$X \text{ und } Y \text{ unabhängig} \implies \varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$$

(dies folgt aus $e^{it(X+Y)} = e^{itX}e^{itY}$ und der Multiplikativität des Erwartungswerts für unabhängige Faktoren).

Weiter ist offenbar für $a, b \in \mathbb{R}$

$$\varphi_{aX+b}(t) = \mathbb{E}\big[e^{it(aX+b)}\big] = e^{itb}\varphi_X(at).$$

Für $Z \sim \mathcal{N}_{0,1}$ ist

$$\varphi_Z(t) = \exp(-t^2/2),$$

ein etwas „hemdsärmeliges“ Argument dafür ist

$$\begin{aligned}\varphi_Z(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-x^2/2}\,dx \\ &= e^{-t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\big(-\tfrac{1}{2}(x-it)^2\big)\,dx = e^{-t^2/2}\end{aligned}$$

(man kann das Integral als die „Dichte einer Normalverteilung mit Erwartungswert it“ interpretieren, die Identität kann z.B. mittels Cauchy'schem Integralsatz bewiesen werden). Siehe auch z.B. [Du, S. 93] für einen „rein reellen“ Beweis.

Bericht 8.2 (Eine Version von Lévys Stetigkeitssatz)**.** $Y, Y_1, Y_2, \dots$ reelle Zufallsvariablen mit

$$\varphi_{Y_n}(t) \underset{n\to\infty}{\longrightarrow} \varphi_Y(t) \quad \text{für jedes } t \in \mathbb{R}.$$

Dann gilt

$$Y_n \underset{n\to\infty}{\longrightarrow} Y$$

(die Y_n konvergieren in Verteilung gegen Y), d.h.

$$\lim_{n\to\infty} \mathbb{P}(Y_n \le y) = \mathbb{P}(Y \le Y) = F_Y(y)$$

für jeden Stetigkeitspunkt von F_Y.

Angesichts von Bericht 8.2 ist ein Weg, Satz 8.1 zu beweisen, die Asymptotik der charakteristischen Funktion von

$$Y_n := \frac{X_1 + \dots + X_n - n\mathbb{E}[X_1]}{\sqrt{n\sigma^2}}$$

zu untersuchen.

Wir können im Folgenden o.E. $\mathbb{E}[X_1] = 0$, $\mathrm{Var}[X_1] = 1$ annehmen, ansonsten betrachten wir

$$\widetilde{X}_i := \frac{X_i - \mathbb{E}[X_1]}{\sqrt{\mathrm{Var}[X_1]}}.$$

Es ist

$$\begin{aligned}\varphi_{Y_n}(t) &= \mathbb{E}\Big[\exp\Big(\frac{it}{\sqrt{n}}(X_1 + \dots + X_n)\Big)\Big] \\ &= \mathbb{E}\Big[\prod_{k=1}^{n} \exp\Big(i\frac{t}{n^{1/2}} X_i\Big)\Big] = \Big(\varphi_{X_1}\big(\tfrac{t}{n^{1/2}}\big)\Big)^n\end{aligned}$$

Taylorentwicklung $e^z = 1 + z + \frac{1}{2}z^2 + o(|z|^2)$ liefert

$$\begin{aligned}\varphi_{X_1}(u) &= \mathbb{E}\big[e^{iuX_1}\big] = 1 + iu\mathbb{E}[X_1] + \frac{i^2u^2}{2}\mathbb{E}[X_1^2] + o(u^2) \\ &= 1 - \frac{u^2}{2} + o(u^2),\end{aligned}$$

also tatsächlich

$$\varphi_{Y_n}(t) = \Big(1 - \frac{t^2}{2n} + o\Big(\frac{t^2}{n}\Big)\Big)^n \underset{n\to\infty}{\longrightarrow} e^{-t^2/2}.$$

8.1 Ein Kopplungsbeweis

Beobachtung. 1. Die Aussage von Satz 8.1 gilt trivialerweise (sogar für festes $n \in \mathbb{N}$), wenn $X_i \sim \mathcal{N}_{0,1}$, da dann

$$\frac{X_1 + \cdots + X_n}{\sqrt{n}} \sim \mathcal{N}_{0,1},$$

Eine Beweisidee für Satz 8.1 ist, den allgemeinen Fall auf diesen Spezialfall zurückzuführen.

2. Wir können o.E. $\mathbb{E}[X_1] = 0$, $\operatorname{Var}[X_1] = 1$ annehmen, ansonsten betrachten wir

$$\widetilde{X}_i := \frac{X_i - \mathbb{E}[X_1]}{\sqrt{\operatorname{Var}[X_1]}}.$$

Lemma 8.3. *$X_1, X_2, \ldots$ u.i.v. reelle ZVn, $X_i \in \mathcal{L}^2$ mit $\mathbb{E}[X_i] = 0$, $\operatorname{Var}[X_i] = 1$, $f : \mathbb{R} \to \mathbb{R}$ dreimal stetig differenzierbar, die ersten drei Ableitungen seien gleichmäßig beschränkt. Dann gilt*

$$\lim_{n\to\infty} \mathbb{E}\Big[f\Big(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\Big)\Big] = \mathbb{E}\big[f(Z)\big]$$

mit $Z \sim \mathcal{N}_{0,1}$.

Beweis. Seien $Z_1, Z_2, \ldots$ u.i.v., $\sim \mathcal{N}_{0,1}$, unabhängig von den X_i, schreibe

$$\begin{aligned} & f\Big(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\Big) - f\Big(\frac{Z_1 + \cdots + Z_n}{\sqrt{n}}\Big) \\ &= \sum_{i=1}^{n} \Big(f\Big(W_{i,n} + \frac{X_i}{\sqrt{n}}\Big) - f\Big(W_{i,n} + \frac{Z_i}{\sqrt{n}}\Big)\Big) \end{aligned} \tag{8.1}$$

mit

$$W_{i,n} := \frac{1}{\sqrt{n}}\Big(X_1 + X_2 + \cdots + X_{i-1} + Z_{i+1} + Z_{i+1} + \cdots + Z_n\Big).$$

Taylor-Enwicklung (im Punkt $W_{i,n}$) liefert

$$\begin{aligned} & f\Big(W_{i,n} + \frac{X_i}{\sqrt{n}}\Big) - f\Big(W_{i,n} + \frac{Z_i}{\sqrt{n}}\Big) \\ &= f'\big(W_{i,n}\big)\frac{X_i - Z_i}{\sqrt{n}} + \frac{1}{2} f''\big(W_{i,n}\big)\frac{X_i^2 - Z_i^2}{n} + R_{i,n} \end{aligned}$$

mit

$$\big|R_{i,n}\big| \le \big|f''\big(W_{i,n} + \vartheta_{i,n}\big) - f''\big(W_{i,n}\big)\big|\frac{X_i^2}{2n} + \big|f''\big(W_{i,n} + \widetilde{\vartheta}_{i,n}\big) - f''\big(W_{i,n}\big)\big|\frac{Z_i^2}{2n}$$

wobei $|\vartheta_{i,n}| \le \dfrac{|X_i|}{\sqrt{n}}$, $|\widetilde{\vartheta}_{i,n}| \le \dfrac{|Z_i|}{\sqrt{n}}$.

Mit $C_2 := \sup_{x\in\mathbb{R}} |f''(x)|$, $C_3 := \sup_{x\in\mathbb{R}} |f'''(x)|$ gilt für jedes $K > 0$:

$$|R_{i,n}| \le C_3 \frac{K^3}{2n^{3/2}} \mathbf{1}_{\{|X_i|\le K\}} + C_2 \frac{X_i^2}{n} \mathbf{1}_{\{|X_i|>K\}} + C_3 \frac{|Z_i|^3}{n^{3/2}}.$$

Nehme Erwartungswert in (8.1):

$$\begin{aligned}
&\Big|\mathbb{E}\Big[f\Big(\frac{X_1+\cdots+X_n}{\sqrt{n}}\Big)\Big] - \mathbb{E}\big[f(Z)\big]\Big| \\
&= \Big|\sum_{i=1}^n \Big(\underbrace{\mathbb{E}\Big[f'\big(W_{i,n}\big)\frac{X_i - Z_i}{\sqrt{n}}\Big]}_{=\mathbb{E}[f'(W_{i,n})]\mathbb{E}[X_i - Z_i]/\sqrt{n}=0} + \underbrace{\mathbb{E}\Big[\frac{1}{2} f''\big(W_{i,n}\big)\frac{X_i^2 - Z_i^2}{n}\Big]}_{=\mathbb{E}[f''(W_{i,n})]\mathbb{E}[X_i^2 - Z_i^2]/(2n)=0} + \mathbb{E}\big[R_{i,n}\big]\Big)\Big| \\
&\le \sum_{i=1}^n \mathbb{E}\big[|R_{i,n}|\big] \le n\Big(C_3 \frac{K^3}{2n^{3/2}} + \frac{C_2}{n}\mathbb{E}\big[X_1^2 \mathbf{1}_{\{|X_1|>K\}}\big] + \frac{C_3}{n^{3/2}}\mathbb{E}\big[|Z_1|^3\big]\Big),
\end{aligned}$$

also

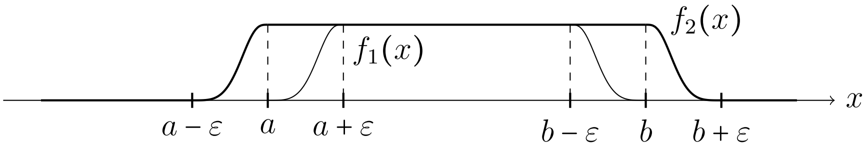
$$\limsup_{n\to\infty} \Big|\mathbb{E}\Big[f\Big(\frac{X_1+\cdots+X_n}{\sqrt{n}}\Big)\Big] - \mathbb{E}\big[f(Z)\big]\Big| \le C_2 \mathbb{E}\big[X_1^2 \mathbf{1}_{\{|X_1|>K\}}\big] \xrightarrow[K\to\infty]{} 0$$

□

Beweis von Satz 8.1. Seien $-\infty < a < b < \infty$. Zu $0 < \varepsilon < \frac{b-a}{2}$ wähle f_1, f_2, die den Voraussetzungen von Lemma 8.3 genügen und

$$\mathbf{1}_{[a+\varepsilon, b-\varepsilon]} \le f_1 \le \mathbf{1}_{[a,b]} \le f_2 \le \mathbf{1}_{[a-\varepsilon, b+\varepsilon]}$$

erfüllen (siehe Skizze).



Es ist

$$\begin{aligned}
\mathbb{P}\big(Z \in [a+\varepsilon, b-\varepsilon]\big) &\le \mathbb{E}\big[f_1(Z)\big] = \lim_{n\to\infty} \mathbb{E}\Big[f_1\Big(\frac{X_1+\cdots+X_n}{\sqrt{n}}\Big)\Big] \\
&\le \liminf_{n\to\infty} \mathbb{P}\Big(\frac{X_1+\cdots+X_n}{\sqrt{n}} \in [a,b]\Big) \\
&\le \limsup_{n\to\infty} \mathbb{P}\Big(\frac{X_1+\cdots+X_n}{\sqrt{n}} \in [a,b]\Big) \\
&\le \lim_{n\to\infty} \mathbb{E}\Big[f_2\Big(\frac{X_1+\cdots+X_n}{\sqrt{n}}\Big)\Big] = \mathbb{E}\big[f_2(Z)\big] \le \mathbb{P}\big(Z \in [a-\varepsilon, b+\varepsilon]\big),
\end{aligned}$$

mit $\varepsilon \downarrow 0$ folgt die Behauptung.

Die Fälle $a = -\infty$ oder $b = +\infty$ kann man analog behandeln. □

Literaturverzeichnis

[Kl] A. Klenke, Wahrscheinlichkeitstheorie, 2. Aufl., Springer, 2008.

[Ka] O. Kallenberg, Foundations of modern probability, 2. ed, Springer, 2002.

[Ge] H.-O. Georgii, Einführung in die Wahrscheinlichkeitstheorie und Statistik, 4. Aufl., de Gruyter, 2009.

[Du] R. Durrett, Probability : theory and examples, Duxbury Press, 2003.

[Wi] D. Williams, Probability with martingales, Cambridge Univ. Press, 1995.

[Fe] W. Feller, An Introduction to Probability Theory, Band 1 und Band 2, Wiley 1968 und 1971.

[Br] L. Breiman: Probability, Wiley, 1968.

[El] J. Elstrodt: Maß- und Integrationstheorie, 5. Auflage, Springer Verlag, 2007.

[Ba] H. Bauer, Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie, de Gruyter, 1978.